# Translation

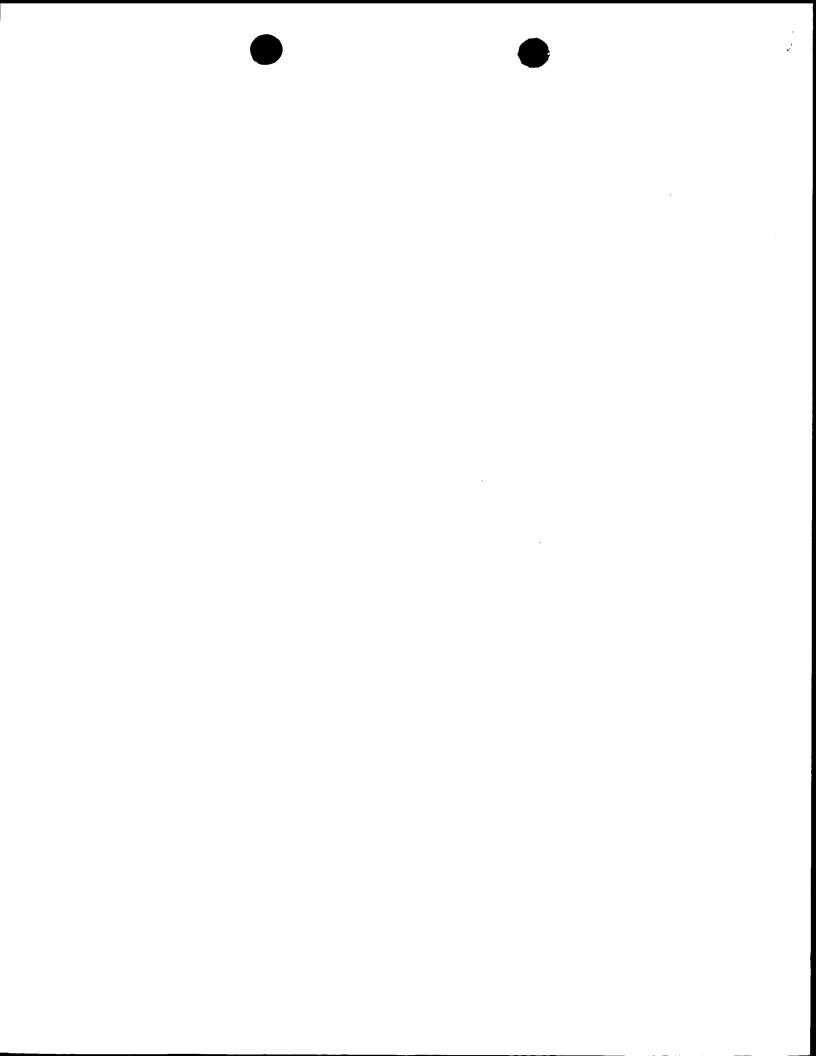
## PATENT COOPERATION TREATY PCT

### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

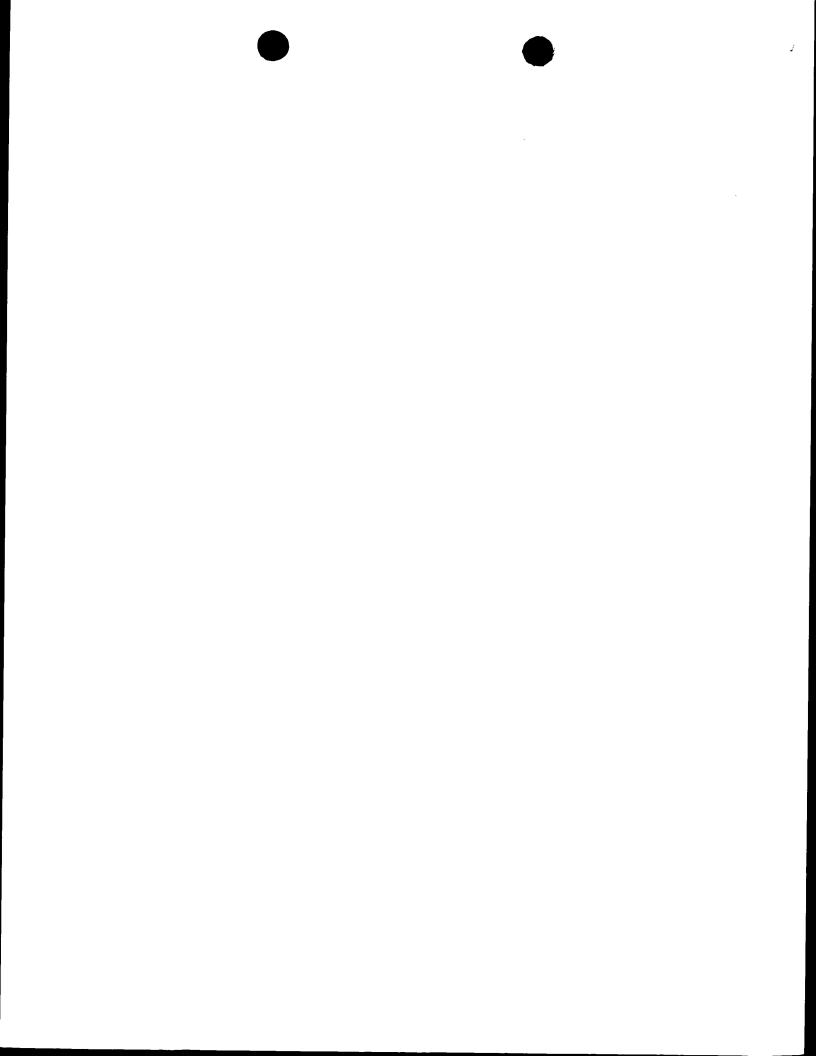
(PCT Article 36 and Rule 70)

7

Applicant's or agent's file reference FZJ 9909 PCT	FOR FURTHER ACTION	Preliminary Ex	tion of Transmittal of International camination Report (Form PCT/IPEA/416)  Priority date (day/month/year)
International application No.  PCT/EP00/07370  International filing of 31 July 200		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	09 August 1999 (09.08.99)
International Patent Classification (IPC) C12N 15/53, Applicant	or national classification and IPC  BASF AKTIENGESEL	LSCHAFT	
Applicant	DASI ARTIERO		
2. This REPORT consists of a tot	tal of 6 sheets, include	ding this cover s	International Preliminary Examining heet.
been amended and are (see Rule 70.16 and S	ompanied by ANNEXES, i.e., sheets the basis for this report and/or sheet ection 607 of the Administrative Instantiation of a total of3 sheets	structions under	tion, claims and/or drawings which have ectifications made before this Authority the PCT).
l .	ns relating to the following items:		
II Priority III Non-establ	ishment of opinion with regard to n	ovelty, inventive	e step and industrial applicability
IV Lack of un	ity of invention statement under Article 35(2) with r nd explanations supporting such sta	regard to novelty tement	y, inventive step or industrial applicability;
1 VI I I	ocuments cited efects in the international application	n	
	bservations on the international appl		
Date of submission of the demand	Ţ		on of this report November 2001 (28.11.2001)
01 March 200  Name and mailing address of the		Authorized offic	eer
Ivanic and manning		Telephone No.	



I. Basis of the report		Office in pageons to an invitation
This report has been drawn of under Article 14 are referred to it.	n the basis of (Replacement sheet: n this report as "originally filed"	s which have been furnished to the receiving Office in response to an invitation and are not annexed to the report since they do not contain amendments.):
the international	application as originally filed.	
the description,	pages1-8	, as originally filed,
	pages	_, filed with the demand,
	pages	
	pages	, nied with the letter of
the claims,	Nos.	, as originally filed,
	Nos.	, as amended under Article 19,
	Nos	
	Nos. 1-20	, filed with the letter of
the drawings,	sheets/fig1-11	, as originally filed,
	sheets/fig	, filed with the letter of,
	sheets/fig	, filed with the letter of
2. The amendments have resu		
	, pages	
the claims,		
-	sheets/fig	_
	1	amendments had not been made, since they have been considered
3. In is report has been to go beyond the dis	closure as filed, as indicated in	the Supplemental Box (Rule 70.2(c)).
4. Additional observations, if	necessary:	
4. Additional observations, is		



International application No. r/EP 00/07370

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

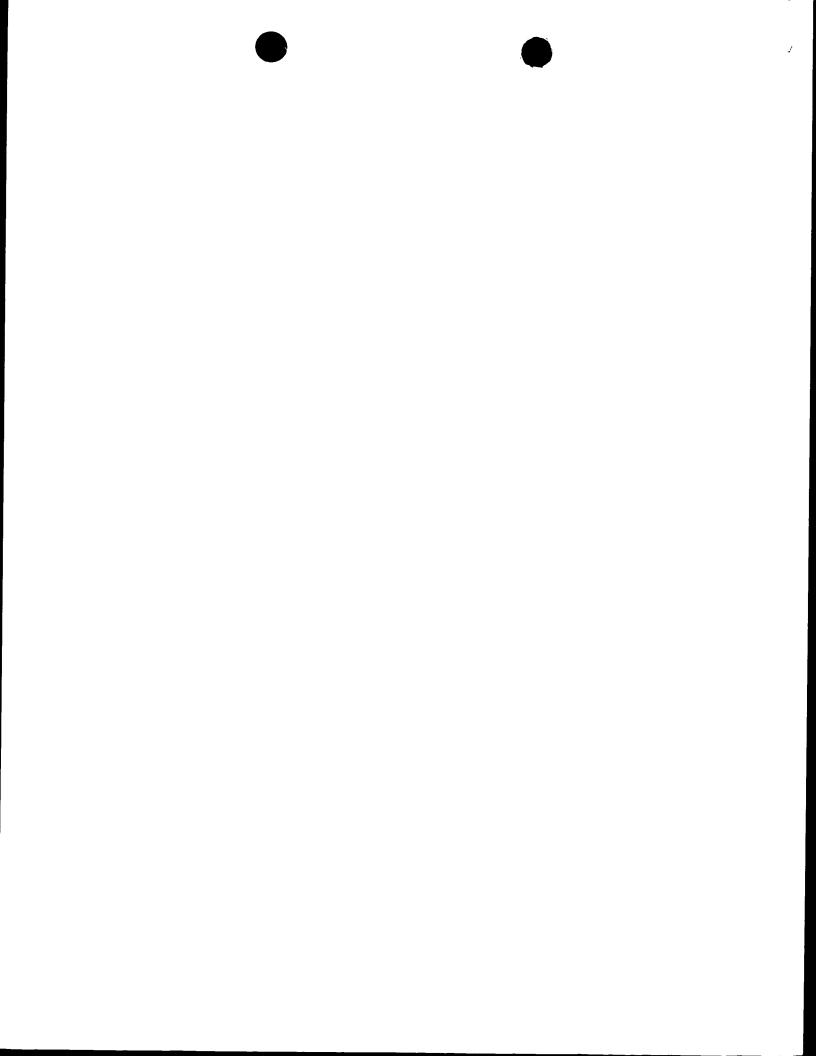
1.	Statement			
	Novelty (N)	Claims	3-5,7,8; 9-14 partially; 15-17;18-20 partially	YES
		Claims	1,2,6; 9-14 partially; 18-20 partially	NO
	Inventive step (IS)	Claims	3-5 partially,7,8; 9-14, partially; 18-20 partially	YES
		Claims	1,2; 3-5 partially; 6; 9-14 partially; 15-17; 18-20 partially	NO

1-20 YES Claims Industrial applicability (IA)

> NO Claims

2. Citations and explanations

SEE SUPPLEMENTAL SHEET



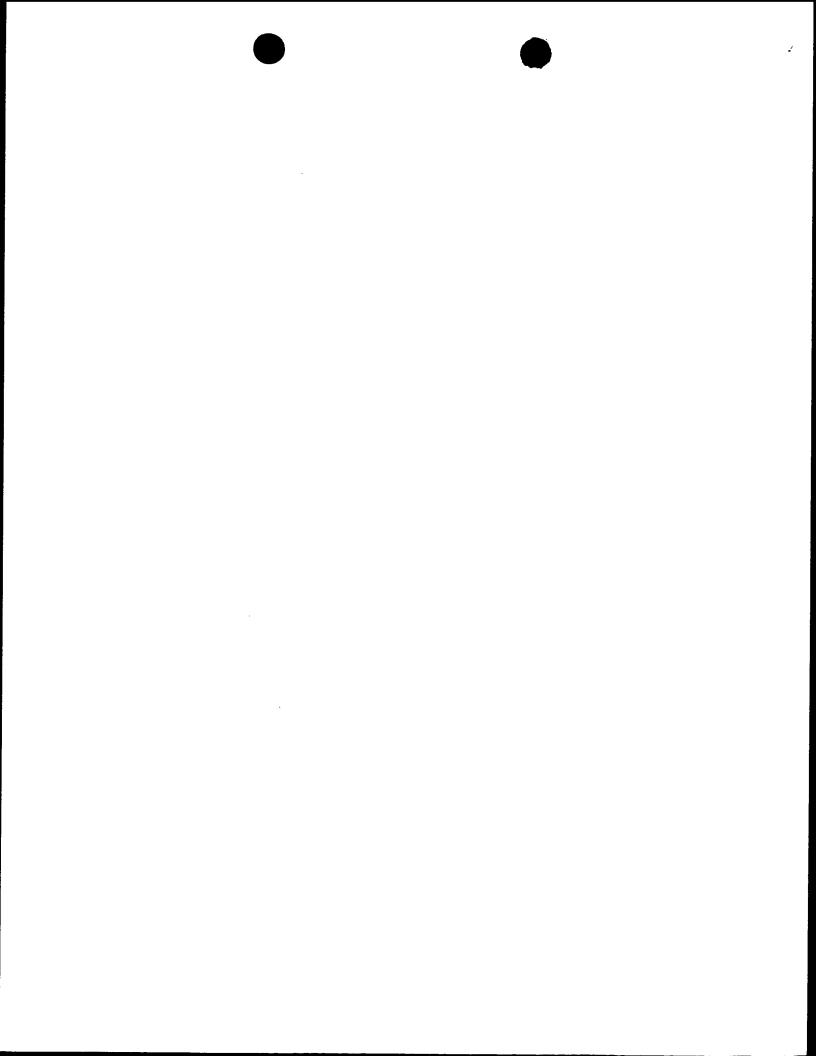
International application No.
PEP 00/07370

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: BOXES I, V and VIII

- The amended Claims 1-20 appear to meet the requirements of PCT Article 34(2)(b). However, it should be noted that the original application documents to which the applicant refers to support the submitted amendments obviously do not match the application documents available to the Examining Authority. The application as originally filed contains an 8-page description. However, the applicant refers for example to a page 11, lines 29-34, with regard to the amended Claims 6, 7 and 8. Likewise, page 7, lines 19-30, does not contain any disclosure relevant to the amended Claim 9.
- PCT Article 33(2) and (3) 2. Part of the amended Claims 1-20 continues not to meet the requirements of PCT Article 33(2) and (3). Document "Henke et al., J. of Biol. Chem., Vol. 273, No. 6, 1998, pages 3702-3711" describes the identification of a peroxisomal isocitrate dehydrogenase (Idp3) and the gene from S. cerevisiae that encodes it, as well as its recombinant expression in E. coli cells transformed accordingly (see Fig. 3, page 3706, left-hand column, second paragraph, and right-hand column, including second paragraph; and page 3709, righthand column, first paragraph). The transformed E. coli cells show a massive increase in isocitrate dehydrogenase activity. Claim 1 concerns a microorganism for the biotechnical production of riboflavin characterised in that it shows higher



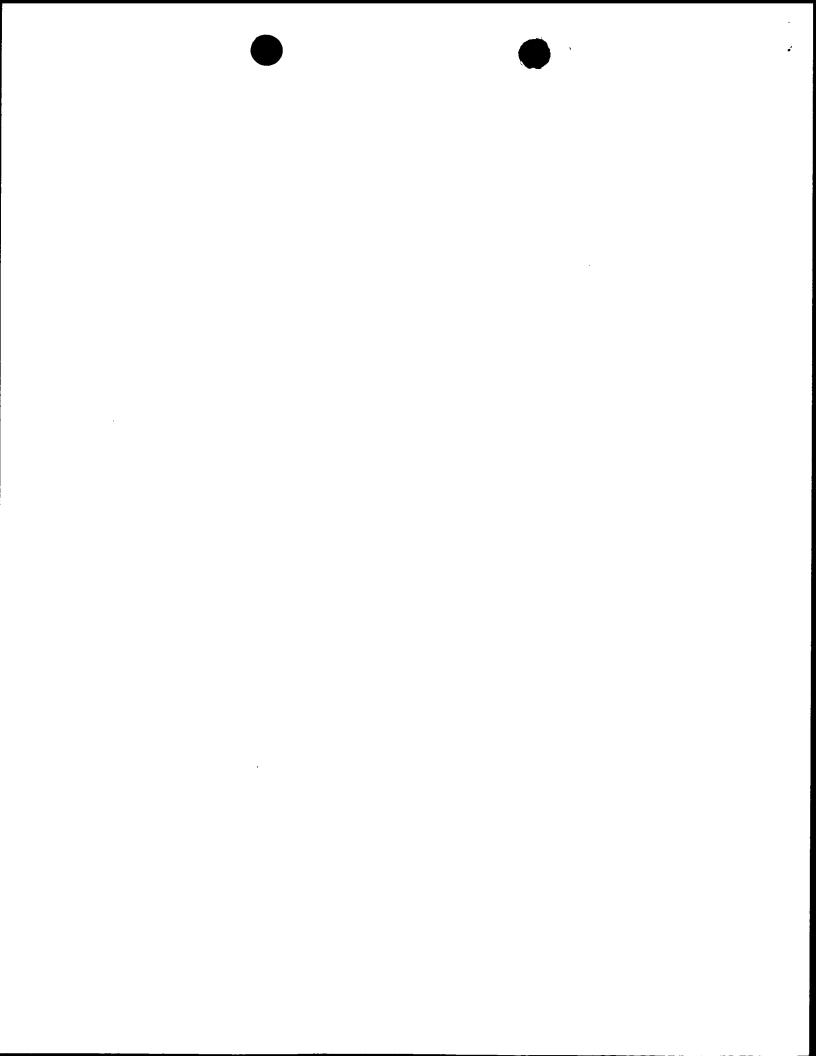
Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: BOXES I, V and VIII

NAD(P)H-forming enzyme activity than a wild type of the species Ashbya gossypii, ATCC 10895 (see also item 3). The intended use, that is the biotechnical production of riboflavin, should be considered to be an optional and hence non-restrictive feature as long as corresponding characterising, specific (i.e. novel) features of the micro-organism which are specially required for this use are not indicated. Consequently, E. coli containing rib genes (see, for example, EP-A-0 405 370, page 3, lines 25-30) would likewise be suitable for the biosynthesis of riboflavin even it this was not demonstrated by Henke et al., since that document highlighted mainly the verification of the cloned Idp3. Moreover, it should be noted that Claim 1 concerns a micro-organism that shows increased activity of an NAD(P)H-forming enzyme. It does not refer to the actual NAD(P)H content produced in the cell. Claim 1 is therefore not considered novel and inventive within the meaning of PCT Article 33(2) and (3). This also applies to dependent Claims 2 and Claims 11, 12, 19 and 20, insofar as they refer directly or indirectly to Claim 6 (see below).

Claim 6 is not considered novel and inventive because the Idp3 gene known from Henke et al. is 65% identical to the DNA sequence of the Idp3 gene described in the present application and is therefore covered by the term "allelic variant" (see the definition of the term "allelic variant" on page 6, lines 3-6). This also applies to Claims



International application No.
EP 00/07370

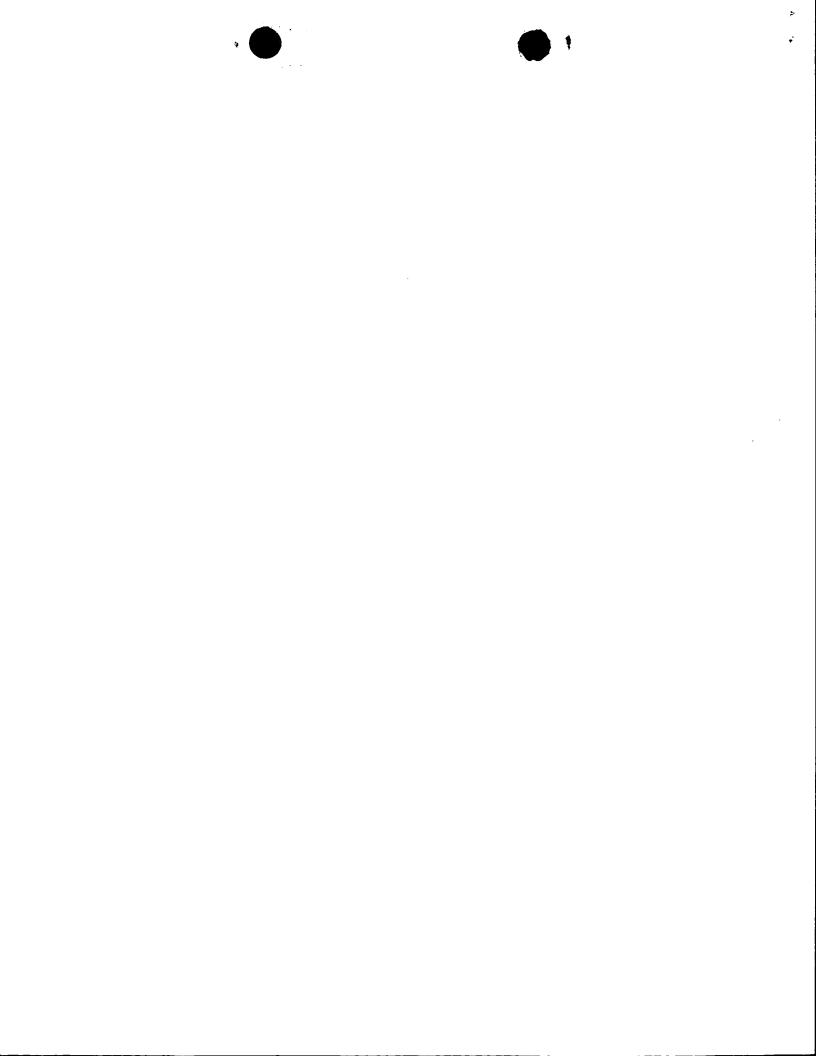
Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: BOXES I, V and VIII

copies of the Idp3 gene, which would both result in an increased activity (see also Claim 11). The circumstances which lead to increased activity are not clear from the present wording. Thus the increased Idp3 activity could be due merely to modified culture conditions. Moreover, a comparison would be meaningful only if two micro-organisms of the same species were compared, such as the Idp3 content in the wild type strain of A. gossypii, ATCC 10895, in specific growth conditions and the Idp3 content of a "genetically modified" strain of A. gossypii, ATCC 10895, in the same growth conditions.

4. Finally, the applicant should note that the PCT and the European proceedings differ from one another, inter alia, in that the last phase of the PCT phase is the establishment of a preliminary examination report, not of a binding decision, and that therefore the right to a legal hearing cannot be harmed.



#### (12) NACH DEM VERTR WEBER DIE INTERNATIONALE ZUSAMMEN ABEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum Internationales Büro



### 

(43) Internationales Veröffentlichungsdatum 15. Februar 2001 (15.02.2001)

**PCT** 

### (10) Internationale Veröffentlichungsnummer WO 01/11052 A3

- (51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/53, 15/80, 9/04, 1/15, C12P 25/00 // (C12N 1/15, C12R 1:645)
- D-52428 Jülich (DE). MAETING, Ines [DE/DE];
   Wiesenstr. 3, D-52428 Jülich (DE).

(21) Internationales Aktenzeichen:

PCT/EP00/07370

(74) Anwalt: FITZNER, Uwe; Lintorfer Strasse 10, D-40878 Ratingen (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,

AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,

HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,

MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,

TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) Internationales Anmeldedatum:

31. Juli 2000 (31.07.2000)

31.07.2000)

(25) Einreichungssprache:

Deutsch

(26) Veröffentlichungssprache:

Deutsch

- (30) Angaben zur Priorität: 199 37 548.8 9. August 1999 (09.08.1999) DE
- (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; D-67056 Ludwigshafen (DE). FORSCHUNGSZEN-TRUM JÜLICH GMBH [DE/DE]; D-52425 Jülich (DE).
- Veröffentlicht:

(72) Erfinder; und

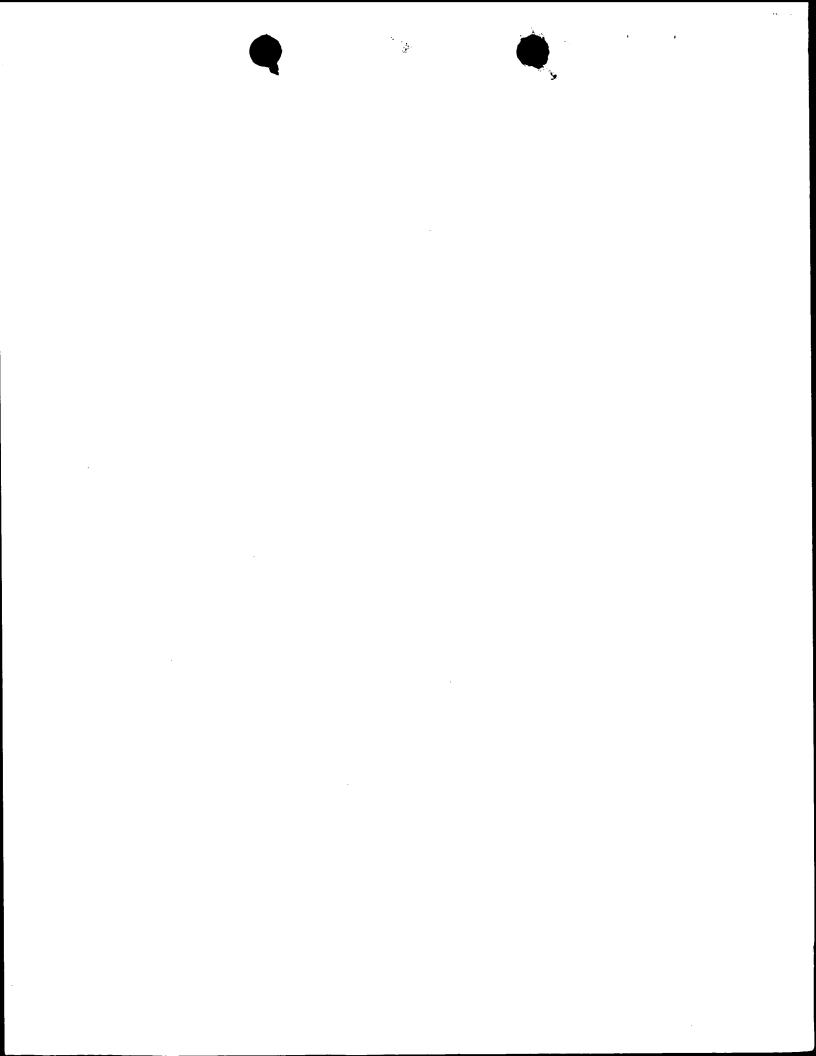
- Mit internationalem Recherchenbericht.
- (75) Erfinder/Anmelder (nur für US): ALTHÖFER, Henning [DE/DE]; Mainstrasse 12, D-67117 Limburgerhof (DE). ZELDER, Oskar [DE/DE]; Rossmarktstrasse 27, D-67346 Speyer (DE). REVUELTA DOVAL, Jose L. [ES/ES]; Grillo 11 4E, E-37001 Salamanca (ES). SANTOS GARCIA, Maria Angeles [ES/ES]; C/Escuelas, 1-5, E-37001 Salamanca (ES). SAHM, Hermann [DE/DE]; Wendelinusstrasse 71, D-52428 Jülich (DE). STAHMANN, Klaus-Peter [DE/DE]; Wilhelmstrasse
- (88) Veröffentlichungsdatum des internationalen Recherchenberichts: 5. Juli 2001

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.



- (54) Bezeichnung: EIN- ODER MEHRZELLIGE ORGANISMEN ZUR HERSTELLUNG VON RIBOFLAVIN
- (57) Abstract: The invention relates to a monocellular or multicellular organism, especially a microorganism, for biotechnological production of riboflavin, whereby the enzymatic activity thereof with respect to NAD(P)H formation is higher than that of a wild type of species Ashbya gossypii ATCC10895.
- (57) Zusammenfassung: Die vorliegende Erfindung betrifft einen ein- oder mehrzelligen Organismus, insbesondere Mikroorganismus, zur biotechnischen Herstellung von Riboflavin, wobei dessen Enzymaktivität bezüglich der NAD(P)H Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.



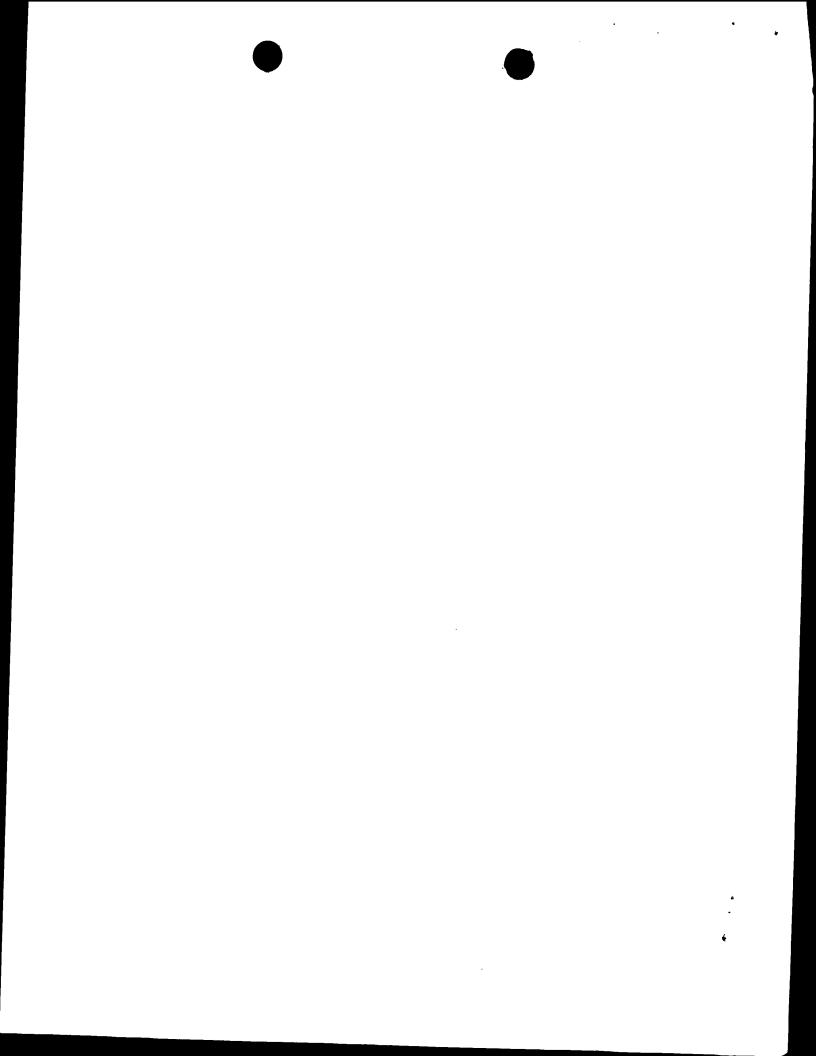


### **PCT**

### INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

(Artikel 36 und Regel 70 PCT)

(Artikel 36 und Reg		
		ung über die Übersendung des internationalen Prüfungsberichts (Formblatt PCT/IPEA/416)
		Prioritätsdatum (Tag/Monat/Tag)
	ag/worlandam/	09/08/1999
31/07/2000		03/03/1000
er nationale Klassifikation und IPK		
et al.		Drifting beauftragten
	•	
amt 6 Blätter einschließlich dies	ses Deckblatts	
ANI ACEN heir dabei handel	t es sich um E	Hätter mit Beschreibungen, Ansprüchen
geändert wurden und diesem Be Berichtigungen (siehe Regel 70.	ericht zugrund 16 und Absch	le liegen, und/oder blatter filt vor dieser nitt 607 der Verwaltungsrichtlinien zum PCT)
	, :	7.4.2
samt 3 Diatter.		
		<del></del>
		<del></del>
ı zu folgenden Punkten:		<del></del> -
n zu folgenden Punkten:		
richts		
richts	erfinderische	Fätigkeit und gewerbliche Anwendbarkeit
richts sines Gutachtens über Neuheit,	erfinderische	Fätigkeit und gewerbliche Anwendbarkeit
richts sines Gutachtens über Neuheit, ( itlichkeit der Erfindung		Fätigkeit und gewerbliche Anwendbarkeit heit, der erfinderischen Tätigkeit und der
richts sines Gutachtens über Neuheit, ( itlichkeit der Erfindung		der orfinderischen Tätigkeit und der
richts eines Gutachtens über Neuheit, d itlichkeit der Erfindung tellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen	htlich der Neu därungen zur	der orfinderischen Tätigkeit und der
richts eines Gutachtens über Neuheit, d itlichkeit der Erfindung tellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen	htlich der Neu därungen zur	der orfinderischen Tätigkeit und der
richts eines Gutachtens über Neuheit, d itlichkeit der Erfindung tellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk	htlich der Neu därungen zur	der orfinderischen Tätigkeit und der
richts eines Gutachtens über Neuheit, d itlichkeit der Erfindung tellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen	htlich der Neu därungen zur	der orfinderischen Tätigkeit und der
richts sines Gutachtens über Neuheit, itlichkeit der Erfindung stellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu klärungen zur g neldung	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung
richts sines Gutachtens über Neuheit, itlichkeit der Erfindung stellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu klärungen zur g neldung	der orfinderischen Tätigkeit und der
richts  sines Gutachtens über Neuheit, ober Meuheit, ober Erfindung stellung nach Artikel 35(2) hinsichtendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu därungen zur g neldung Datum der Ferlig	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung
richts  sines Gutachtens über Neuheit, ober Meuheit, ober Erfindung stellung nach Artikel 35(2) hinsichtendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu klärungen zur g neldung	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung
richts  sines Gutachtens über Neuheit, ditlichkeit der Erfindung  stellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu därungen zur g neldung Datum der Ferlig 28.11.2001	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung 
richts  sines Gutachtens über Neuheit, ditlichkeit der Erfindung  stellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu därungen zur g neldung Datum der Ferlig	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung  gstellung dieses Berichts
richts  sines Gutachtens über Neuheit, of  itlichkeit der Erfindung  itellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk  ührte Unterlagen  el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu därungen zur g neldung Datum der Ferlig 28.11.2001	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung  gstellung dieses Berichts
richts  sines Gutachtens über Neuheit, of itlichkeit der Erfindung itellung nach Artikel 35(2) hinsicl vendbarkeit; Unterlagen und Erk ührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann	htlich der Neu därungen zur g neldung  Datum der Ferlig 28.11.2001  Bevollmächtigte	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung gstellung dieses Berichts r Bediensteter
sines Gutachtens über Neuheit, of itlichkeit der Erfindung stellung nach Artikel 35(2) hinsich vendbarkeit; Unterlagen und Erkührte Unterlagen el der internationalen Anmeldung rkungen zur internationalen Ann stemationalen vorläufigen	htlich der Neu därungen zur g neldung Datum der Ferlig 28.11.2001	heit, der erfinderischen Tätigkeit und der Stützung dieser Feststellung gstellung dieses Berichts r Bediensteter
	amt 6 Blätter einschließlich dies	WEITERES VORGEHEN vorläufigen Internationales Anmeldedatum(Tag/Monat/Jahr) 31/07/2000 er nationale Klassifikation und IPK  et al.  Prüfungsbericht wurde von der mit der internationale gemäß Artikel 36 übermittelt.  amt 6 Blätter einschließlich dieses Deckblatts cht ANLAGEN bei; dabei handelt es sich um Egeändert wurden und diesem Bericht zugrund Berichtigungen (siehe Regel 70.16 und Absch



### INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

Internationales Aktenzeichen PCT/EP00/0737

۱.	Grundlage des Berichts							
1.	Hinsichtlich der Bestandteile der internationalen Anmeldung (Ersatzblätter, die dem Anmeldeamt auf eine Aufforderung nach Artikel 14 hin vorgelegt wurden, gelten im Rahmen dieses Berichts als "ursprünglich eingereicht" und sind ihm nicht beigefügt, weil sie keine Änderungen enthalten (Regeln 70.16 und 70.17)): Beschreibung, Seiten:							
1-8 ursprüngliche Fassung								
	Pate	entansprüche, Nr	::					
	1-20	)	eingegangen am	13/09/2001	mit Schreiben vom	10/09/2001		
	Zei	chnungen, Blätter	r:					
	1-1	1	ursprüngliche Fassung					
2. Hinsichtlich der <b>Sprache</b> : Alle vorstehend genannten Bestandteile standen der Behörde in der Sprache, i die internationale Anmeldung eingereicht worden ist, zur Verfügung oder wurden in dieser eingereicht, so unter diesem Punkt nichts anderes angegeben ist.						in der Sprache, in der er eingereicht, sofern		
Die Bestandteile standen der Behörde in der Sprache: zur Verfügung bzw. wurden in dieser Sprache eingereicht; dabei handelt es sich um								
		die Sprache der ( Regel 23.1(b)).	Übersetzung, die für die Zwe	cke der internatio	nalen Recherche ein	gereicht worden ist (nach		
		die Veröffentlichu	ingssprache der internationa	len Anmeldung (n	ach Regel 48.3(b)).			
			Übersetzung, die für die Zwe 5.2 und/oder 55.3).	cke der internatio	nalen vorläufigen Prü	fung eingereicht worden		
3. Hinsichtlich der in der internationalen Anmeldung offenbarten <b>Nucleotid- und/oder Aminosäuresequenz</b> is internationale vorläufige Prüfung auf der Grundlage des Sequenzprotokolls durchgeführt worden, das:					nosäuresequenz ist die worden, das:			
	☐ in der internationalen Anmeldung in schriftlicher Form enthalten ist.							
			er internationalen Anmeldung			worden ist.		

4. Aufgrund der Änderungen sind folgende Unterlagen fortgefallen:

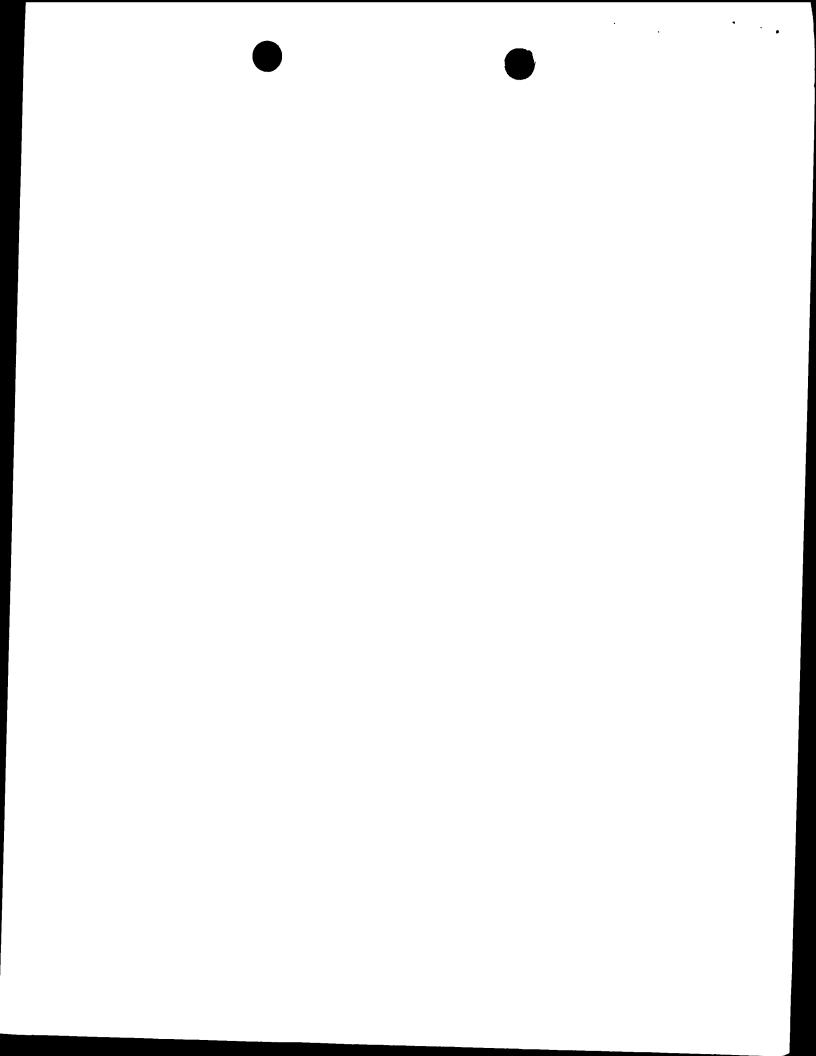
Sequenzprotokoll entsprechen, wurde vorgelegt.

☐ bei der Behörde nachträglich in computerlesbarer Form eingereicht worden ist.

☐ Die Erklärung, daß das nachträglich eingereichte schriftliche Sequenzprotokoll nicht über den

Die Erklärung, daß die in computerlesbarer Form erfassten Informationen dem schriftlichen

Offenbarungsgehalt der internationalen Anmeldung im Anmeldezeitpunkt hinausgeht, wurde vorgelegt.



### INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

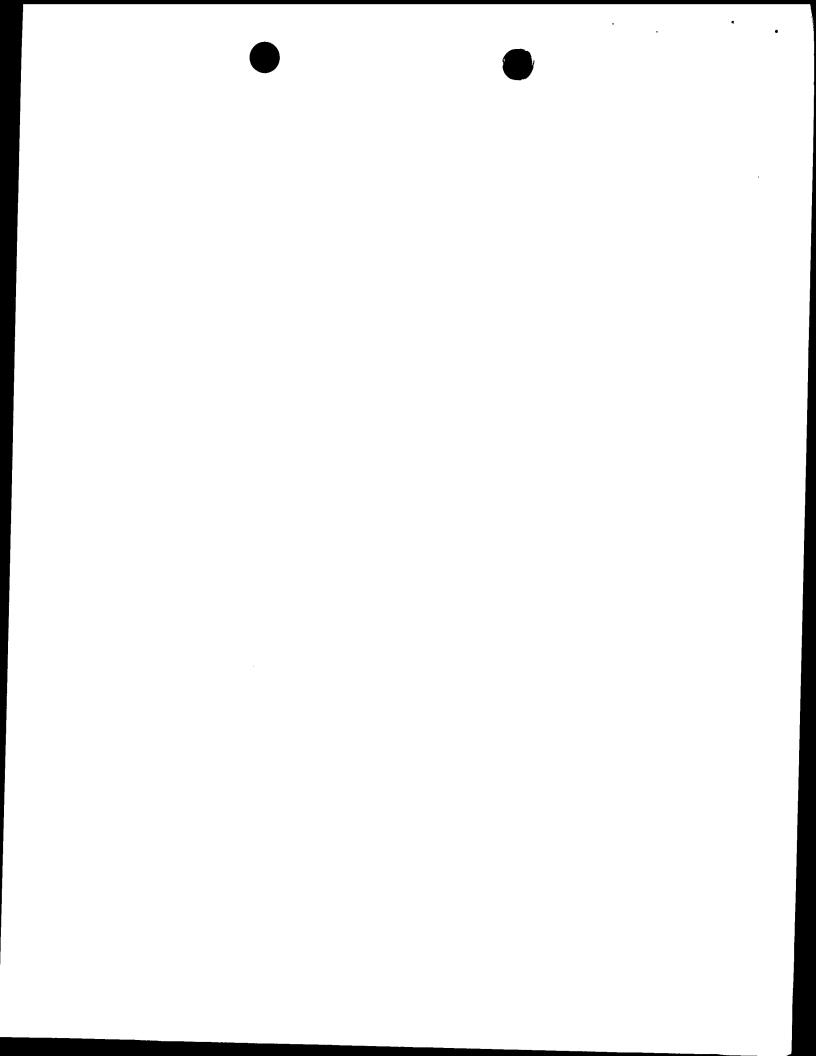
Internationales Aktenzeichen PCT/EP00/07370

		Beschreibung,	Seiten:			
		Ansprüche,	Nr.:			
		Zeichnungen,	Blatt:			
		•	-			
5.	□ Dieser Bericht ist ohne Berücksichtigung (von einigen) der Änderungen erstellt worden, da diese aus den angegebenen Gründen nach Auffassung der Behörde über den Offenbarungsgehalt in der ursprünglich eingereichten Fassung hinausgehen (Regel 70.2(c)).					
			loha Ändari	una	en enthalten	ist unter Punkt 1 hinzuweisen;sie sind diesem Bericht
		(Auf Ersatzblätter, di beizufügen).	e soiche Anderd	ung	en entration,	
6.	Etw	vaige zusätzliche Bem	nerkungen:			•
٧	<ul> <li>Begründete Feststellung nach Artikel 35(2) hinsichtlich der Neuheit, der erfinderischen T\u00e4tigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erkl\u00e4rungen zur St\u00fctzung dieser Feststellung</li> </ul>					
1	. Fe	ststellung				
	Ne	euheit (N)	Ja: Ne		Ansprüche Ansprüche	3-5, 7, 8; 9-14 teilweise; 15-17; 18-20 teilweise 1, 2, 6; 9-14, teilweise; 18-20 teilweise
	Eri	finderische Tätigkeit (l	ET) Ja Ne		Ansprüche Ansprüche	3-5 teilweise, 7, 8; 9-14, teilweise; 18-20 teilweise 1, 2; 3-5 teilweise; 6; 9-14 teilweise; 15-17; 18-20 teilweise
	Ge	ewerbliche Anwendba	rkeit (GA) Ja Ne		Ansprüche Ansprüche	1-20

2. Unterlagen und Erklärungen siehe Beiblatt

### VIII. Bestimmte Bemerkungen zur internationalen Anmeldung

Zur Klarheit der Patentansprüche, der Beschreibung und der Zeichnungen oder zu der Frage, ob die Ansprüche in vollem Umfang durch die Beschreibung gestützt werden, ist folgendes zu bemerken: siehe Beiblatt

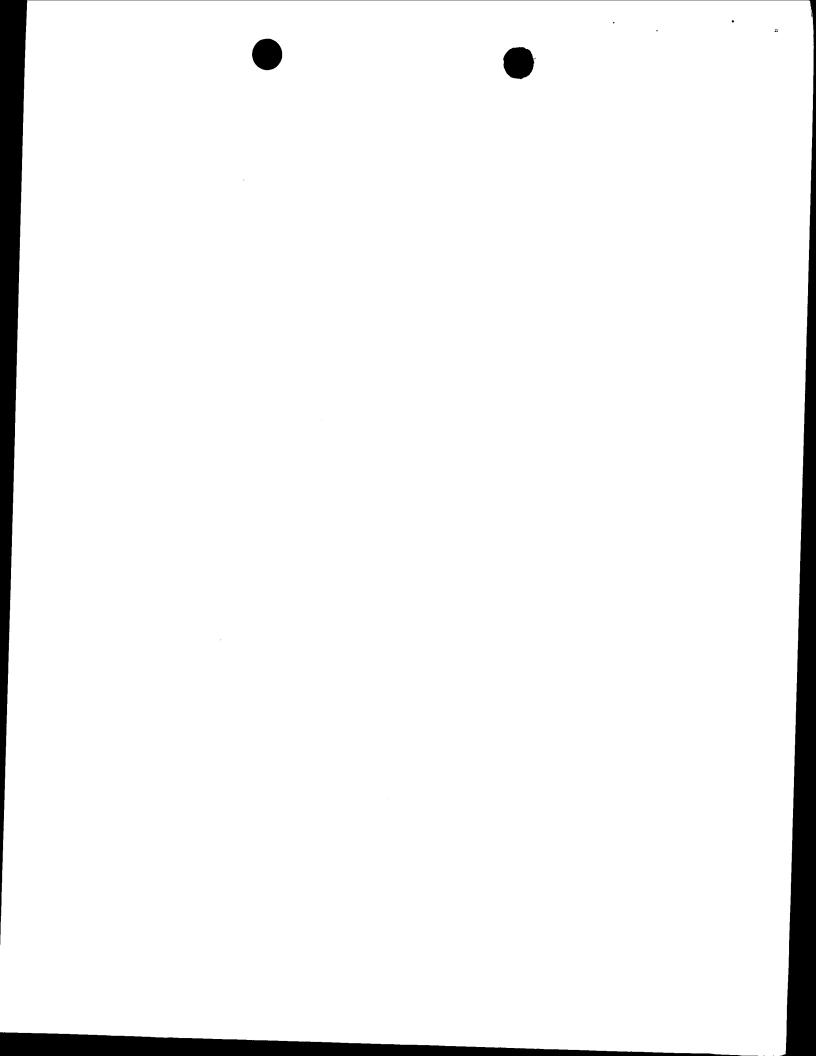


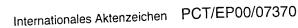
### 1. Art. 34 (2,b) PCT

Die geänderten Ansprüche 1-20 scheinen den Anforderungen des Art. 34(2,b) PCT zu entsprechen. Es sei in diesem Zusammenhang aber darauf hingewiesen, dass offensichtlich die ursprünglichen Anmeldungsunterlagen, auf die zur Stützung der eingeführten Änderungen Bezug genommen wurde, nicht mit den der Behörde vorliegenden Anmeldungsunterlagen übereinstimmen. Die Anmeldung wie ursprünglich eingereicht, enthält eine Beschreibung von acht Seiten. Mit Hinblick auf die geänderten Ansprüche 6, 7 und 8 wurde aber beispielsweise auf eine Seite 11, Zeilen 29-34 Bezug genommen. Ebenso befindet sich auf Seite 7, Zeilen 19-30 keine sinngemässe Offenbarung für den geänderten Anspruch 9.

### 2. Art. 33(2), (3) PCT

Ein Teil der geänderten Ansprüche 1-20 scheint nach wie vor nicht den Anforderungen des Art. 33(2), (3) PCT zu entsprechen. Dokument "Henke et al., J. of Biol Chem., Bd. 273, Nr. 6, 1998, 3702-3711", beschreibt die Identifizierung einer peroxisomalen Isocitrat-Dehydrogenase (Idp3) und des sie kodierenden Gens aus S. cerevisiae, sowie dessen rekombinante Expression in entsprechend transformierten E. coli Zellen (s. Fig. 3, S. 3706, linke Spalte, zweiter Absatz und rechte Spalte, einschl. zweiter Absatz, Seite 3709, rechte Spalte erster Absatz). Die transformierten E. coli Zellen zeigten einen massiven Anstieg der Isocitratedehydrogenase-Aktivität. Anspruch 1 bezieht sich auf einen Mikroorganismus zur biotechnischen Herstellung von Riboflavin, der dadurch gekennzeichnet ist, dass er eine Aktivität eines NAD(P)H bildenden Enzyms aufweist, die höher ist als diejenige eines Wildtyps der Species Ashbya gossypii ATCC 10895 (s. auch unter Punkt 3.). Der beabsichtigte Verwendungszweck, also die biotechnische Herstellung von Riboflavin, ist als optionales und daher nicht limitierendes Merkmal anzusehen, solange nicht den Mikroorganismus entsprechend kennzeichnende spezifische (d.h. neuheitsgebende) und speziell für diesen Zweck benötigte Merkmale aufgeführt sind. Daher wäre E. coli, welches rib Gene enthält (s. beispielsweise E0 0 405 370, Seite 3, Zeilen 25-30) gleichermassen für die Biosynthese von Riboflavin geeignet, auch wenn dies von Henke et al. nicht nachgewiesen wurde, da hier lediglich die Verifizierung der klonierten Idp3 im Vordergrund stand. Zudem sei darauf hingewiesen, dass sich Anspruch 1 auf einen Mikroorganismus bezieht, der eine erhöhte Aktivität







LIKE THERE SERVICES IN

bezüglich eines NAD(P)H bildenden Enzyms aufweist. Zur Höhe des tatsächlich in der Zelle gebildeten NAD(P)H-Gehalts wird auch hier nicht Bezug genommen. Anspruch 1 wird daher nicht als neu und erfinderisch im Sinne von Art. 33(2), (3) PCT angesehen. Dies gilt ebenso für den abhängigen Anspruch 2 und die Ansprüche 11, 12, 19 und 20, soweit sie sich direkt oder indirekt auf Anspruch 6 rückbeziehen (s. unten).

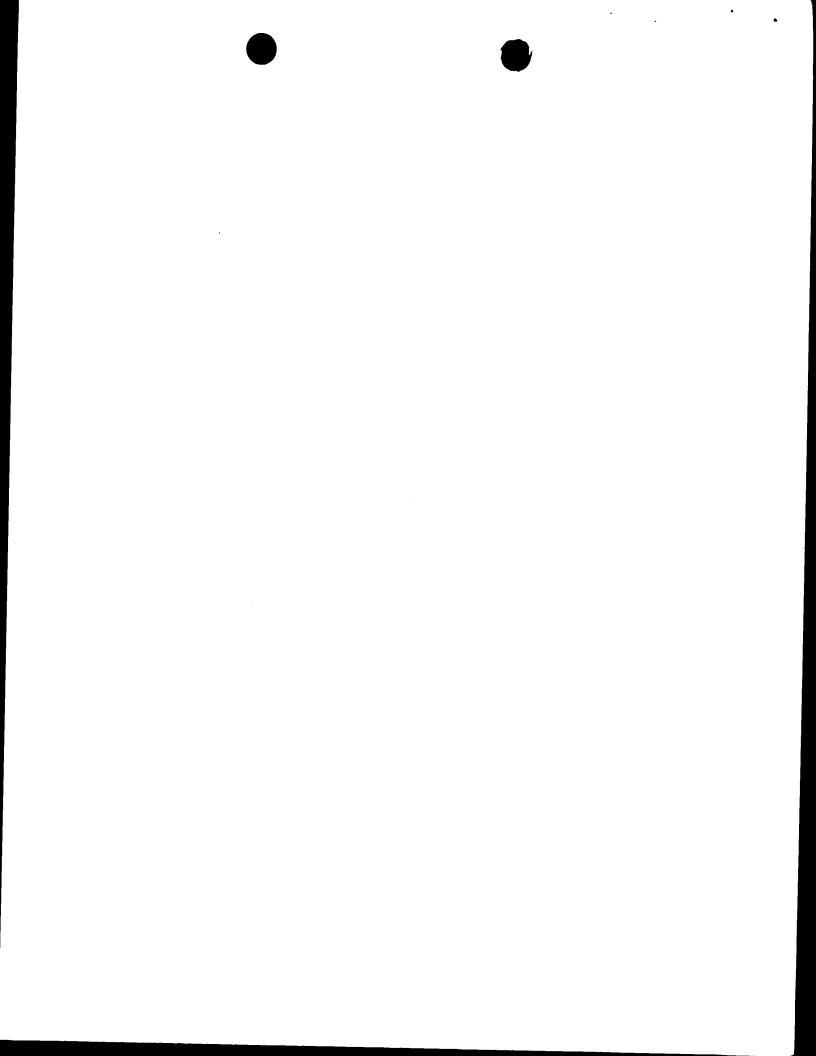
Anspruch 6 wird nicht als neu und erfinderisch angesehen, da das aus Henke et al. bekannte ldp3-Gen zu 65% identisch ist mit der in der gegenwärtigen Anmeldung beschriebenen DNA-Sequenz des Idp3-Gens und somit unter den Begriff "Allelvariation" fällt (s. Definition des Begriffes Allelvariation auf Seite 6, Zeilen 3-6). Dies gilt ebenfalls für die Ansprüche 9 und 10.

Die Ansprüche 3-5 scheinen sich auf einen neuen und soweit sie sich auf Anspruch 2 rückbeziehen auch auf einen erfinderischen Gegenstand zu beziehen. Dies gilt auch für die direkt abhängigen Ansprüche 14, 18. Ansprüche 7, 8 sowie davon direkt oder indirekt abhängige Ansprüche 9-14 und 18 können ebenfalls als neu und erfinderisch angesehen werden.

#### Art. 5/6 PCT

Der gegenwärtigen Erfindung liegt die Entdeckung zugrunde, dass mit einer verstärkten Isocitrat-dehydrogenase-Expression im Pilz Ashbya ATCC 10895 auch ein deutlicher Anstieg der Riboflavinbildung einhergeht. Allerdings lässt sich aus dieser Beobachtung kein allgemein gültiges Prinzip ableiten, wonach der Anstieg irgendeines anderen NAD(P)H bildenden Enzyms zum gleichen Ergebnis führt. Mit Hinblick auf die auf Idp3 limitierte Beschreibung, sind die Ansprüche, die lediglich vage auf ein NAD(P)H bildendes Enzym gerichtet sind, zu breit und daher nicht mit den Erfordernissen des Art. 5 vereinbar.

Desweiteren sei darauf hingewiesen, dass ohne spezifischen Vergleichswert oder klare Definition des Vergleichswerts (bspw. ldp3-Gehalt unter spezifischen Wachstumbedingungen) oder Angabe eines präzisen, strukturellen Merkmals (bspw. verstärkte Expression oder erhöhte Kopienzahl des Idp3-Gens, die beide eine erhöhte Aktivität zur Folge haben, s. auch Anspruch 11), Anspruch 1 (auch Ansprüche 2-3) unklar ist (Art. 6 PCT), da aus der derzeitigen Formulierung nicht eindeutig hervorgeht, aufgrund welcher Umstände die Aktivität erhöht ist. So könnten auch lediglich geänderte Kulturbedingungen eine erhöhte Idp3-Aktivität bewirken. Zudem scheint ein Vergleich nur zwischen zwei Mikroorganismen der



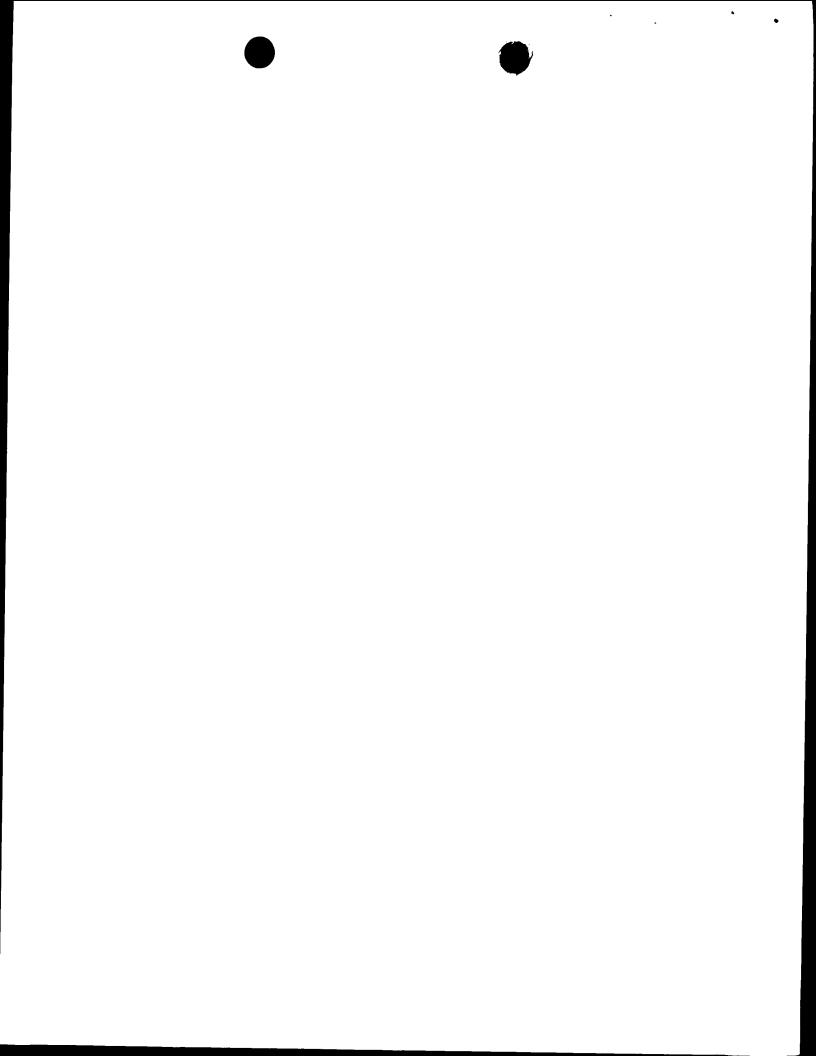


### INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT - BEIBLATT

Internationales Aktenzeichen PCT/EP00/07370

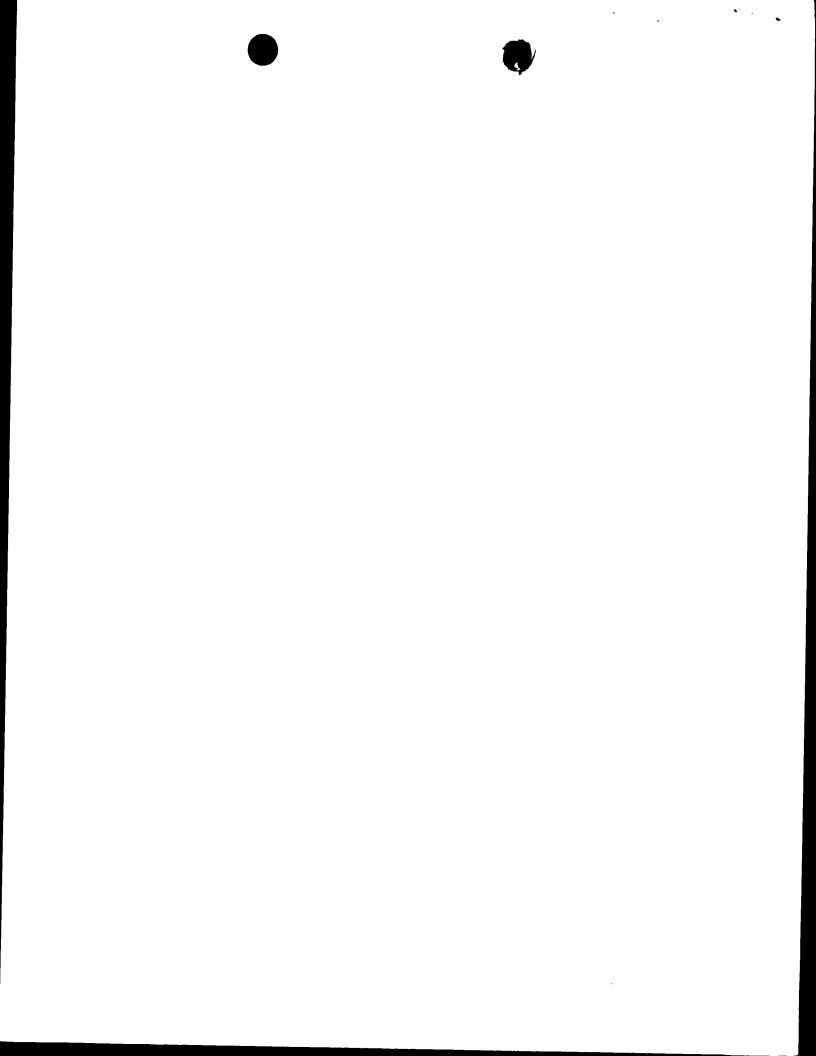
gleichen Species sinnvoll, d.h. Idp3-Gehalt im Wildtypstamm A. gossypii ATCC10895 unter entsprechend spezifizierten Wachstumsbedingungen und Idp3-Gehalt eines "genetisch veränderten" Stamms von A. gossypii ATCC10895 unter den gleichen Wachstumsbedingungen.

Schliesslich sei der Anmelder darauf hingewiesen, dass sich das PCT und Euro-4. Verfahren unter anderem darin unterscheiden, dass in der PCT-Phase als letzte Aktion lediglich ein vorläufiger Prüfungsbericht erstellt wird, aber keine bindende Entscheidung getroffen wird, sodass ein Recht auf rechtliches Gehör nicht verletzt werden kann.

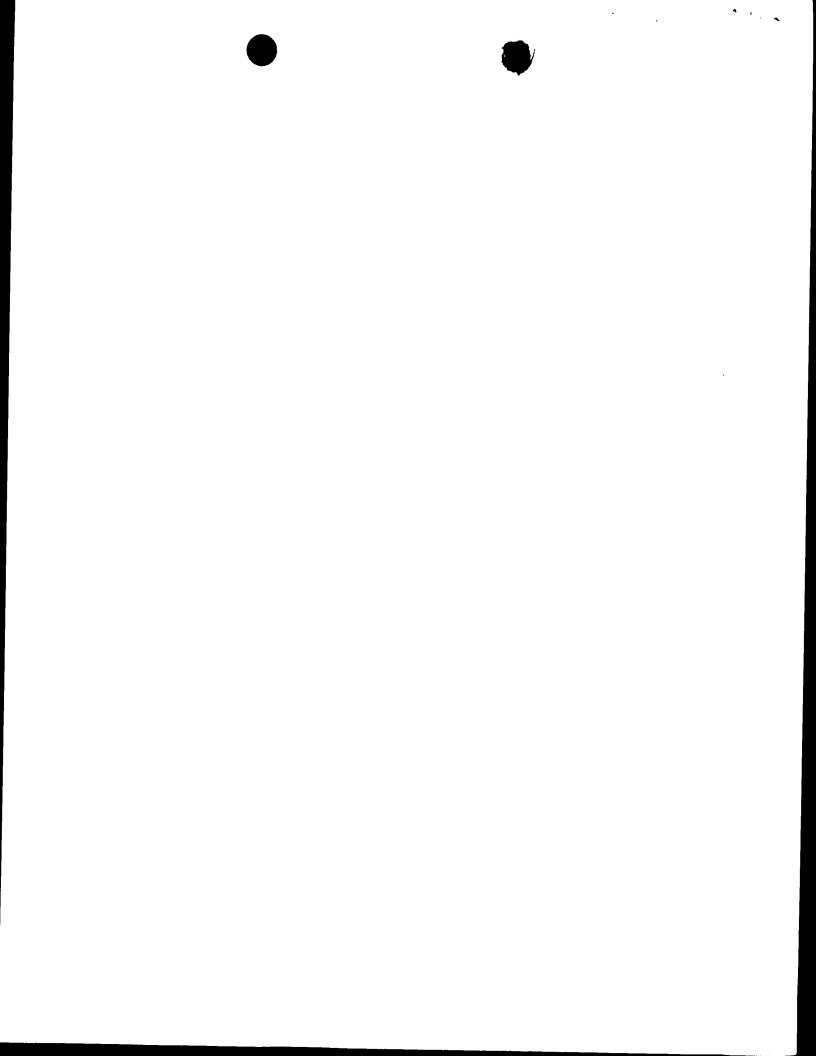


#### Neue Ansprüche 1-20:

- Ein- oder mehrzelliger Mikroorganismus zur biotechnischen Herstellung von Riboflavin, dadurch gekennzeichnet, daß er eine Aktivität eines NAD(P)H-bildenden. Enzyms aufweist, die h\u00f6her ist als diejenige eines Wildtyps der Species Ashbya gossypii ATCC10895.
- Ein- oder mehrzelliger Mikroorganismus nach Anspruch 1,
   dadurch gekennzeichnet, daß er eine erhöhte Isocitrat-Dehydrogenase-Aktivität aufweist.
- Ein- oder mehrzelliger Mikroorganismus nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß er ein Pilz ist.
- Ein- oder mehrzelliger Mikroorganismus nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß er ein Pilz aus der Gattung Ashbya ist.
- Ein- oder mehrzelliger Mikroorganismus nach einem der Ansprüche 1 bis 4,
   dadurch gekennzeichnet, daß er ein Pilz der Spezies Ashbya gossypii ist.
- 6. Isocitrat-Dehydrogenase-Gen mit einer für die in Fig. 11 (SEQ ID No. 2) angegebenen Aminosäuresequenz und deren Allelvariationen kodierenden Nukleotidsequenz.
- 7. Isocitrat-Dehydrogenase-Gen nach Anspruch 6 mit der Nukleotidsequenz von Nukleotid 1 bis 1262 gem. der Fig. 11 (SEQ ID No. 1).
- Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 oder 7 mit einem vorgeschalteten Promotor der Nukleotidsequenz mit Nukleotid –661 bis –1 gem. der Fig. 11 (SEQ ID No. 1).
- Gen-Struktur enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche
   bis 8 sowie mit diesem operativ verknüpfte regulatorische Sequenzen.
- Vektor enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 8 oder eine Gen-Struktur nach Ansprüch 9.

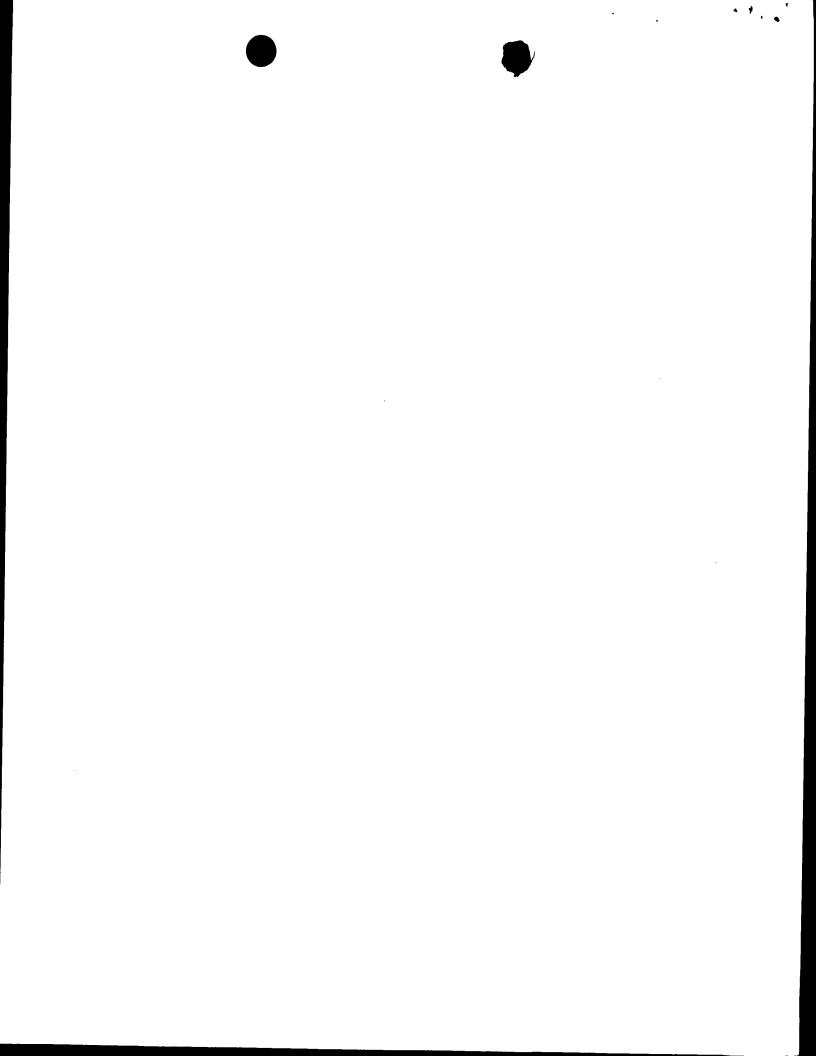


- 11. Genetisch veränderter Mikroorganismus zur biotechnischen Herstellung von Riboflavin enthaltend in replizierbarer Form ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 8, welches im Vergleich zu dem entsprechend nicht genetisch veränderten Mikroorganismus verstärkt exprimiert wird und/oder deren Kopienzahl erhöht ist.
- 12. Genetisch veränderter Mikroorganismus gemäß Anspruch 11 enthaltend eine Gen-Struktur nach Anspruch 9 oder einen Vektor nach Anspruch 10.
- 13. Genetisch veränderter Mikroorganismus nach einem der Ansprüche 11 oder 12 enthaltend eine Isocitrat-Dehydrogenase, die im Vergleich zu dem entsprechend nicht genetisch veränderten Mikroorganismus eine gesteigerte katalytische Aktivität und/oder eine verringerte Hemmbarkeit gegenüber Inhibitoren aufweist.
- 14. Verfahren zur biotechnischen Herstellung von Riboflavin, dadurch gekennzeichnet, daß ein Mikroorganismus gemäß einem der Ansprüche 1 bis 5 oder 11 bis 13 eingesetzt wird.
- 15. Verfahren zur Herstellung eines Riboflavin produzierenden ein- oder mehrzelligen Mikroorganismus, dadurch gekennzeichnet, daß durch gentechnische Methoden die Aktivität eines NAD(P)H-bildenden Enzyms im Vergleich zu derjenigen eines Wildtyps der Species Ashbya gossypii ATCC10895 erhöht wird.
- 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß die Erhöhung der Enzymaktivität durch Austausch des Promotors und/oder Erhöhung der Genkopienzahl erzielt wird.
- 17. Verfahren nach einem der Ansprüche 15 oder 16, dadurch gekennzeichnet, daß die Erhöhung der Enzymaktivität durch eine gesteigerte katalytische Aktivität und/oder eine verringerte Hemmbarkeit gegenüber Inhibitoren der Isocitrat-Dehydrogenase erzeugt wird.
- 18. Verwendung eines Mikroorganismus nach einem der Ansprüche 1 bis 5 sowie 11 bis 13 zur Herstellung von Riboflavin.





- 19. Verwendung eines Isocitrat-Dehydrogenase-Gens nach einem der Ansprüche 6 bis 8 zur Herstellung eines Mikroorganismus nach einem der Ansprüche 1 bis 5 sowie 11 bis 13.
- Verwendung einer Gen-Struktur nach Anspruch 9 oder eines Vektors nach Anspruch
   10 zur Herstellung eines Mikroorganismus nach einem der Ansprüche 1 bis 5 sowie
   11 bis 13.



### F ENT COOPERATION TREAT

#### **PCT**

### **NOTIFICATION OF ELECTION**

(PCT Rule 61.2)

### From the INTERNATIONAL BUREAU

l To:

Commissioner
US Department of Commerce
United States Patent and Trademark
Office, PCT
2011 South Clark Place Room
CP2/5C24
Arlington, VA 22202

Date of mailing (day/month/year)

12 April 2001 (12.04.01)

ETATS-UNIS D'AMERIQUE
in its capacity as elected Office

International application No. PCT/EP00/07370

Applicant's or agent's file reference FZJ 9909 PCT

International filing date (day/month/year) 31 July 2000 (31.07.00)

Priority date (day/month/year)
09 August 1999 (09.08.99)

Applicant

ALTHÖFER, Henning et al

1.	The designated Office is hereby notified of its election made:							
	X in the demand filed with the International Preliminary Examining Authority on:							
	01 March 2001 (01.03.01)							
	in a notice effecting later election filed with the International Bureau on:							
2.	The election X was							
	was not							
	made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).							

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland

Authorized officer

Zakaria EL KHODARY

Facsimile No.: (41-22) 740.14.35

Telephone No.: (41-22) 338.83.38

	·	

FZJ 9909\_PCT zentrum Jülich GmbH BASF Amengesellschaft

### Ein- oder mehrzellige Organismen zur Herstellung von Riboflavin 5

Die vorliegende Erfindung betrifft einen ein- oder mehrzelligen Organismus zur Herstellung von Riboflavin.

Das Vitamin B2, auch Riboflavin genannt, ist für Mensch und Tier 10 essentiell. Bei Vitamin-B2-Mangel treten Entzündungen der Mund- und den Mundwinkeln, Juckreiz und Rachenschleimhäute, Risse in Entzündungen in den Hautfalten u. a. Hautschäden, Bindehautentzündungen, verminderte Sehschärfe und Trübung der Hornhaut auf. Bei Säuglingen und Kindern können Wachstumsstillstand und Gewichtsabnahme eintreten. Das Vitamin B2 hat daher wirtschaftliche Bedeutung 15 insbesondere als Vitaminpräparat bei Vitaminmangel sowie als Futtermittelzusatz. Daneben wird es auch als Lebensmittelfarbstoff, beispielsweise in Mayonnaise, Eiscreme, Pudding etc., eingesetzt.

20

30

Die Herstellung von Riboflavin erfolgt entweder chemisch oder mikrobiell. Bei den chemischen Herstellungsverfahren wird das Riboflavin in der Regel in mehrstufigen Prozessen als reines Endprodukt gewonnen, wobei allerdings auch relativ kostspielige Ausgangsprodukte - wie beispielsweise D-Ribose - eingesetzt werden müssen.

25

Eine Alternative zur chemischen Herstellung des Riboflavins bietet die Herstellung dieses Stoffes durch Mikroorganismen. Die mikrobielle Herstellung des Riboflavins eignet sich insbesondere für solche Fälle, in denen eine hohe Reinheit dieser Substanz nicht erforderlich ist. Dies ist beispielsweise dann der Fall, wenn das Riboflavin als Zusatz zu Futtermittelprodukten eingesetzt werden soll. In solchen Fällen hat die mikrobielle Herstellung des Riboflavins den Vorteil, daß diese Substanz in können Auch ist. gewinnbar einstufigen Prozeß einem

5 Ausgangsprodukte für die mikrobielle Synthese nachwachsende Rohstoffe, wie beispielsweise pflanzliche Öle, eingesetzt werden.

FZJ 9909\_PCT

20

25

30

35

Die Herstellung von Riboflavin durch Fermentation von Pilzen wie Ashbya gossypii oder Eremothecium ashbyi ist bekannt (The Merck Index, Windholz et al., eds. Merck & Co., Seite 1183, 1983, A. Bacher, F. Lingens, Augen. Chem. 1969, S. 393); aber auch Hefen, wie z.B. Candida oder Saccharomyces, und Bakterien wie Clostridium, Bacillus und Corynebakterium sind zur Riboflavinproduktion geeignet.

Zudem sind Verfahren mit der Hefe Candida famata beispielsweise in der US 05231007 beschrieben.

Riboflavin-überproduzierende Bakterienstämme sind beispielsweise in der EP 405370 beschrieben, wobei die Stämme durch Transformation der Riboflavin- Biosynthese-Gene aus Bacillus subtilis erhalten wurden. Diese Prokaryonten-Gene waren aber für ein rekombinantes Riboflavin-Herstellungsverfahren mit Eukaryonten wie Saccharomyces cerevisiae oder Ashbya gossypii ungeeignet. Daher wurden gemäß der WO 93/03183 für die Riboflavin-Biosynthese spezifische Gene aus einem Eukaryonten, nämlich aus Saccharomyces cerevisiae, isoliert, um damit ein rekombinantes Herstellungsverfahren für Riboflavin in einem Derartige bereitzustellen. Produktionsorganismus eukaryontischen rekombinante Herstellungsverfahren haben für die Riboflavin-Produktion jedoch dann keinen oder nur begrenzten Erfolg, wenn die Bereitstellung von Substrat für die an der Riboflavin-Biosynthese spezifisch beteiligten Enzyme unzureichend ist.

1967 fand Hanson (Hanson AM, 1967, in Microbial Technology, Peppler, HJ, pp.222-250 New York), daß der Zusatz der Aminosäure Glycin die Riboflavin-Bildung von Ashbya gossypii steigert. Ein derartiges Verfahren

ist jedoch nachteilig, weil Glycin ein sehr teurer Rohstoff ist. Aus diesem Grunde war man bestrebt, durch Herstellung von Mutanten die Riboflavin-Produktion zu optimieren.

Aus der DE 19545468.5 A1 ist ein weiteres Verfahren zur mikrobiellen

Herstellung von Riboflavin bekannt, bei dem die Isocitratlyase-Aktivität oder die Isocitratlyase-Genexpression eines Riboflavin produzierenden Mikroorganismus erhöht ist. Darüber hinaus ist aus der DE 19840709 A1 ein ein- oder mehrzelliger Organismus insbesondere ein Mikroorganismus zur biotechnischen Herstellung von Riboflavin bekannt. Dieser zeichnet sich dadurch aus, daß einen derart veränderten Glycinstoffwechsel aufweist, daß seine Riboflavinsyntheseleistung ohne externe Zufuhr von Glycin mindestens gleich derjenigen eines Wildtyps der Species Ashbya gossypii ATCC10892 ist.

20 Aber auch im Vergleich zu diesen Verfahren besteht noch ein Bedarf zu einer weiteren Optimierung der Riboflavin-Herstellung.

Aufgabe der vorliegenden Erfindung ist es demgemäß, einen ein- oder mehrzelligen Organismus, vorzugsweise einen Mikroorganismus, für die biotechnische Herstellung von Riboflavin zur Verfügung zu stellen, der eine weitere Optimierung der Riboflavin-Bildung ermöglicht. Insbesondere sollte ein Organismus zur Verfügung gestellt werden, der eine Produktion ermöglicht, die gegenüber dem bisherigen Stand der Technik wirtschaftlicher ist. Vor allem soll der Organismus eine im Vergleich zu den bisherigen Organismen erhöhte Riboflavin-Bildung erlauben.

Diese Aufgabe wird durch einen ein- oder mehrzelligen Organismus gelöst, dessen Enzymaktivität der bezüglich NAD(P)H-Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.

30

25

FZJ 9909 PCT

20

25

30

35

Das Ziel einer beschleunigten intrazellulären NAD(P)H-Versorgung kann durch Erhöhung der Aktivität eines NAD(P)H-bildenden bzw. Senkung der Aktivität eines NAD(P)H verbrauchenden Enzyms bzw. eine Änderung der Spezifität erreicht werden. Dies läßt sich mit den bekannten Methoden der Stammverbesserung von Organismen erreicht werden. D. h. im einfachsten Falle lassen sich entsprechende Stämme nach der in der Mikrobiologie üblichen Selektion mittels Screening herstellen. Ebenso ist die Mutation mit anschließender Selektion einsetzbar. Die Mutation kann hierbei sowohl mittels chemischer als auch mittels physikalischer Mutagenese ausgeführt werden. Eine weitere Methode ist die Selektion und Mutation mit anschließender Rekombination. Schließlich lassen sich die erfindungsgemäßen Organismen mittels Genmanipulation herstellen.

Erfindungsgemäß wird der Organismus derart verändert, daß er intrazellulär NAD(P)H in einer Menge erzeugt, die größer als sein Bedarf für die Aufrechterhaltung seines Metabolismus ist. Diese Erhöhung der erfindungsgemäß sich NAD(P)H-Erzeugung läßt intrazellulären vorzugsweise dadurch erreichen, daß ein Organismus hergestellt wird, bei dem die Enzymaktivität der Isocitrat-Dehydrogenase erhöht ist. Dies kann beispielsweise dadurch erreicht werden, daß durch Veränderung des katalytischen Zentrums ein erhöhter Substratumsatz erfolgt oder indem die Wirkung von Enzyminhibitoren aufgehoben wird. Auch kann eine erhöhte Enzymaktivität der Isocitrat-Dehydrogenase durch Erhöhung der Enzymsynthese, beispielsweise durch Genamplifikation oder durch Ausschaltung von Faktoren, die die Enzym-Biosynthese reprimieren, hervorgerufen werden.

Die Isocitrat-Dehydrogenase-Aktivität kann erfindungsgemäß vorzugsweise durch Mutation des Isocitrat-Dehydrogenase-Gens erhöht werden. Derartige Mutationen können entweder nach klassischen Methoden ungerichtet erzeugt werden, wie beispielsweise durch UV-

Bestrahlung oder mutationsauslösende Chemikalien, oder gezielt mittels gentechnologischer Methoden, wie Deletion, Insertion und/oder Nukleotid-Austausch.

Die Isocitrat-Dehydrogenase-Genexpression kann durch Einbau von Isocitrat-Dehydrogenase-Genkopien und/oder durch Verstärkung regulatorischer Faktoren, die die Isocitrat-Dehydrogenase-Genexpression positiv beeinflussen, erreicht werden. So kann eine Verstärkung regulatorischer Elemente vorzugsweise auf Transcriptionsebene erfolgen, indem insbesondere die Transcriptionssignale erhöht werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der m-RNA verbessert wird.

Zur Erhöhung der Genkopienzahl kann beispielsweise das Isocitrat-Dehydrogenase-Gen in ein Genkonstrukt bzw. in einen Vektor eingebaut werden, der vorzugsweise dem Isocitrat-Dehydrogenase-Gen zugeordnete regulatorische Gensequenzen enthält, insbesondere solche, die die Genexpression verstärken. Anschließend wird ein Riboflavin-produzierender Mikroorganismus, mit dem das Isocitrat-Dehydrogenase-Gen enthaltenden Genkonstrukt transformiert.

25

30

20

Erfindungsgemäß kann die Überexpression der Isocitrat-Dehydrogenase auch durch Austausch des Promotors erzielt werden. Hierbei ist es möglich, die höhere enzymatische Aktivität alternativ durch Einbau von Genkopien oder durch Austausch des Promotors zu erzielen. Gleichermaßen ist es jedoch auch möglich, durch gleichzeitigen Austausch des Promotors und Einbau von Genkopien die gewünschte Änderung der Enzymaktivität zu erzielen.

Die Veränderung des Isocitrat-Dehydrogenase-Gens führt zu einer 35 beschleunigten NAD(P)H-Bildung und zugleich zu einer überraschend

20

25

30

5 hohen Steigerung der Riboflavin-Bildung, wie sie bisher nicht erreichbar war.

Das Isocitrat-Dehydrogenase-Gen wird vorzugsweise aus Mikroorganismen, besonders bevorzugt aus Pilzen, isoliert. Dabei sind Pilze der Gattung Ashbya wiederum bevorzugt. Höchst bevorzugt ist die Spezies Ashbya gossypii.

Für die Isolierung des Gens kommen aber auch alle weiteren Organismen, deren Zellen die Sequenz zur Bildung der Isocitrat-Dehydrogenase enthalten, also auch pflanzliche und tierische Zellen, in Betracht. Die durch homologe oder heterologe Isolierung des Gens kann Komplementation einer im Isocitrat-Dehydrogenase-Gen defekten Mutante oder auch durch heterologes Probing oder PCR mit heterologen Primern erfolgen. Zur Subklonierung kann das Insert des komplementierenden Plasmids anschließend durch geeignete Schritte mit Restriktionsenzymen in der Größe minimiert werden. Nach Sequenzierung und Identifizierung des putativen Gens erfolgt eine paßgenaue Subklonierung durch PCR. Plasmide, die die so erhaltenen Fragmente als Insert tragen, werden in die Isocitrat-Dehydrogenase-Gen-Defekte Mutante eingebracht, die auf Isocitrat-Dehydrogenase-Gens getestet Funktionalität des Funktionelle Konstrukte werden schließlich zur Transformation eines Riboflavin-Produzenten eingesetzt.

Nach Isolierung und Sequenzierung sind die Isocitrat-Dehydrogenase-Gene mit Nukleotidsequenzen erhältlich, die für die angegebene Aminosäure-Sequenz oder deren Allelvariation kodieren. Allelvariationen umfassen insbesondere Derivate, die durch Deletion, Insertion und Substitution von Nuldeotiden aus entsprechenden Sequenzen erhältlich sind, wobei die Isocitrat-Dehydrogenase-Aktivität erhalten bleibt. Eine

15

30

35

5 entsprechende Sequenz ist in Figur 2b von Nukleotid 1 bis 1262 angegeben.

Den Isocitrat-Dehydrogenase-Genen ist insbesondere ein Promotor der Nukleotidsequenz von Nukleotid –661 bis –1 gem. Fig. 11 oder eine im wesentlichen gleich wirkende DNA-Sequenz vorgeschaltet. So kann beispielsweise dem Gen ein Promotor vorgeschaltet sein, der sich von dem Promotor mit der angegebenen Nukleotidsequenz durch ein oder mehrere Nukleotidaustausche, durch Insertion und/oder Deletion unterscheidet, ohne daß aber die Funktionalität bzw. die Wirksamkeit des Promotors beeinträchtigt wird. Des weiteren kann der Promotor durch Veränderung seiner Sequenz in seiner Wirksamkeit erhöht oder komplett durch wirksame Promotoren ausgetauscht werden.

Dem Isocitrat-Dehydrogenase-Gen können des weiteren regulatorische Gen-Sequenzen bzw. Regulatorgene zugeordnet sein, die insbesondere die Isocitrat-Dehydrogenase-Gen-Aktivität erhöhen. So können dem Isocitrat-Dehydrogenase-Gen beispielsweise sog. "enhancer" zugeordnet sein, die über eine verbesserte Wechselwirkung zwischen RNA-Polymerase und DNA eine erhöhte Isocitrat-Dehydrogenase-Expression bewirken.

Dem Isocitrat-Dehydrogenase-Gen mit oder ohne vorgeschaltetem Promotor bzw. mit oder ohne Regulator-Gen können ein oder mehrere DNA-Sequenzen vor- und/oder nachgeschaltet sein, so daß das Gen in einer Gen-Struktur enthalten ist. Durch Klonierung des Isocitrat-Dehydrogenase-Gens sind Plasmide bzw. Vektoren erhältlich, die das Isocitrat-Dehydrogenase-Gen enthalten und zur Transformation eines Riboflavin-Produzenten geeignet sind. Die durch Transformation erhältlichen Zellen enthalten das Gen in replizierbarer Form, d.h. in zusätzlichen Kopien auf dem Chromosom, wobei die Genkopien durch



homologe Rekombination an beliebigen Stellen des Genoms integriert werden und/oder auf einem Plasmid bzw. Vektor.

Bei den erfindungsgemäß erhaltenen ein- oder mehrzelligen Organismen kann es sich um beliebige für biotechnische Verfahren einsetzbare Zellen handeln. Hierzu zählen beispielsweise Pilze, Hefen, Bakterien sowie pflanzliche und tierische Zellen. Erfindungsgemäß handelt es sich vorzugsweise um transformierte Zellen von Pilzen, besonders bevorzugt von Pilzen der Gattung Ashbya. Hierbei ist die Spezies Ashbya gossypii besonders bevorzugt.

15

35

10

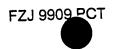
Im folgenden wird die Erfindung näher anhand von Beispielen erläutert, ohne daß damit eine Begrenzung auf den Gegenstand der Beispiele verbunden sein soll:

Das Gen der Isocitrat-Dehydrogenase (IDP3) wurde durch PCR kloniert und dann sequenziert (Sequenz siehe Fig. 11). Die gentechnisch durchgeführte partielle Deletion des Gens durch Austauschmutagenese mit einem Geneticinresistenz-Gen (Fig. 1) wurde durch Southem Blot (Fig. 2) bestätigt. Diese Disruption, d.h. Zerstörung des Gens im Genom des Pilzes, führt dazu, daß der Pilz die davon kodierte Isocitrat-Dehydrogenase nicht mehr bilden kann. Fig. 3 zeigt die Abnahme der Enzymaktivität im Disruptionsstamm AgΔDP3b im Vergleich zum Wildtyp ATCC 10895. In Präparationen der Peroxisomen konnte gezeigt werden, daß dieses Enzym in diesen Organellen lokalisiert ist (Fig. 10). Während die Enzymaktivität in Wildtyp-Peroxisomen deutlich messbar ist, findet sich in den Peroxisomen des Disruptionsstamms keine Aktivität mehr.

Die Disruption des Gens führt zu einer deutlichen Verminderung der Vitaminbildung im Vergleich zum Elternstamm (Fig. 4). Wird das Gen dagegen unter Steuerung des starken TEF-Promotors auf einem Plasmid

(Fig. 6) in zusätzlicher Kopie in die Ashbya-Zellen gebracht, ist ein deutlicher Anstieg in der Enzymaktivität und der Riboflavinbildung meßbar (Fig. 5).

Fig. 7 zeigt, daß bei der Verstoffwechselung ungesättigter Fettsäuren NADPH bei zwei von drei alternativen Reaktionswegen als Reduktionsmittel benötigt wird. Die darin involvierte 2,4-Dienoyl-CoA-Reduktase konnte in Zellen von Ashbya ebenfalls in Peroxisomen lokalisiert werden (Fig. 8). Die Disruption des IDP3-Gens sollte nun zu einem verringerten Wachstum der Zellen auf Linolsäure oder Linolensäure führen. Das konnte auch gemessen werden (Fig. 9). Damit zeigt sich, daß die Bedeutung der IDP3 für den Stoffwechsel der Zelle in der NADPH-Bildung liegt.



## 5 Beschreibung der Figuren

- Fig. 1: Schema der Konstruktion des Vektors pIDPkan für den Genaustausch des chromosomalen *AgIDP*3-Gens gegen eine durch Deletion und Insertion des G418<sup>R</sup>-Gens inaktive Genkopie.
- 10 Fig. 2: Überprüfung der partiellen Deletion und gleichzeitigen Insertion der Geneticin-Resistenz-Kassette am *AgIDP*-Lokus mittels Southern-Blot-Analyse. Genomische, *Sph*I-gespaltene DNA wurde mit einer Digoxygenin-markierten Sonde hybridisiert.
- 15 Fig. 3: Vergleich der Enzymaktivitäten der NADP-spezifischen ICDH vom Ashbya-Wildtyp, der Mutante A.g. ΔIDP3b und den AgIDP-Überexprimierern A.g. pAGIDP3a und A.g. pAGIDP3b bei Wachstum auf Glucose-Vollmedium.

: .

- Fig. 4: Vergleich des Wachstums und der Riboflavinbildung vom 20 Ashbya-Wildtyp und der Mutante A.g. ΔIDP3b bei Wachstum auf Sojaöl-Vollmedium.
- Fig. 5: Vergleich des Wachstums, der Riboflavinbildung und der NADP-spezifischen ICDH vom *Ashbya*-Wildtyp und den *AgIDP*3-Überexprimierern *A.g.* pAGIDP3a und *A.g.* pAGIDP3b bei Kultivierung auf Sojaöl-Vollmedium.
  - Fig. 6: Plasmid zur Überexpression des *AgIDP3*-Gens unter Kontrolle von *TEF*-Promotor und *TEF*-Terminator.

    Zur Einführung der *Sph*I-Schnittstelle war eine Änderung der für die zweite Aminosäure kodierenden Nukleotidsequenz notwendig. Es wurde ein konservativer Austausch der Aminosäure Glycin in Leucin vorgenommen.

15

11
Abbauwege ungesättigter Fettsäuren mit Doppelbindungen

- Fig. 7: Abbauwege ungesättigter Fettsäuren mit Doppelbindungen an geraden (A) und ungeraden (B, C) C-Atomen in Peroxisomen nach Henke et al. (1998).
- 10 Fig. 8 Trennung von aus Ashbya-Wildtyp isolierten Organellen im Percoll-Dichtegradienten:

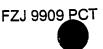
Aktivitäten [U/ml] der Markerenzyme Katalase (Peroxisomen) und Fumarase (Mitochondrien), der NAD- und der NADP- spezifischen ICDH und der für den Abbau ungesättigter Fettsäuren notwendigen 2,4-Dienoyl-CoA-Reduktase und  $\Delta^3,\Delta^2$ - Enoyl-CoA-Isomerase

- Fig. 9: Vergleich des radialen Wachstums von Asbya-Wildtyp, der Mutanten A.g. ΔIDP3a und A.g. ΔIDP3b und den Überexprimierern A.g. pAGIDP3a und A.g. pAGIDP3b auf verschiedene Fettsäuren (A: 18:1 cis9; b: 18:2 cis9,12; C: 18:3 cis9,12,15).
- Fig. 10: Verteilung der Enzyme Katalase und ICDH im Percoll-Dichtegradienten nach Zentrifugation von Organellen aus Myzel des Wildtyps (A) und der Mutante *A.g.* Δ*IDP*3b (B).

# Beschreibung der Sequenz

30

Nukleotidsequenz und daraus abgeleitete Aminosäuresequenz des für die peroxisomale NADP-spezifische Isocitrat-Dehydrogenase codierenden *AgIDP3*-Gens aus *A. gossypii*.



#### **Ansprüche** 5

- Ein- oder mehrzelliger Organismus, insbesondere Mikroorganismus, 1. zur biotechnischen Herstellung von Riboflavin, dadurch gekennzeichnet, daß dessen Enzymaktivität bezüglich der NAD(P)H-Bildung höher ist als diejenige eines Wildtyps der 10 Species Ashbya gossypii ATCC10895.
- Ein- oder mehrzelliger Organismus nach Anspruch 1, 2. dadurch gekennzeichnet, daß er eine erhöhte Isocitrat-Dehydrogenase-Aktivität aufweist. 15
  - Ein- oder mehrzelliger Organismus nach einem der Ansprüche 1 3. oder 2, dadurch gekennzeichnet, daß er ein Pilz ist.

20 Ein- oder mehrzelliger Organismus nach einem der Ansprüche 1 bis 4. 3, dadurch gekennzeichnet, daß er ein Pilz aus der Gattung Ashbya ist.

Ein- oder mehrzelliger Organismus nach einem der Ansprüche 1 bis 5. 4, dadurch gekennzeichnet, daß er ein Pilz der Spezies Ashbya gossypii ist.

Isotitrat-Dehydrogenase-Gen mit einer für die 6. Allelvariation angegebenen Aminosäuresequenz und deren kodierenden Nukleotidsequenz.

25

30

- Isocitrat-Dehydrogenase-Gen nach Anspruch 6 mit der Nukleotidsequenz von Nukleotid 1 bis 1262 gem. der Fig. 11 oder einer im wesentlichen gleich wirkenden DNA-Sequenz.
- Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 oder 7
   mit einem vorgeschalteten Promotor der Nukleotidsequenz mit Nukleotid –661 bis –1 gem. der Fig. 11 oder einer im wesentlichen gleich wirkenden DNA-Sequenz.
- Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 8 mit
   diesem zugeordneten regulatorischen Gensequenzen.
  - 10. Gen-Struktur enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 9.
- 20 11. Vektor enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 9 oder eine Gen-Struktur nach Anspruch 10.
- 12. Transformierter Organismus zur Herstellung von Riboflavin enthaltend in replizierbarer Form ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 9 oder eine Gen-Struktur nach Anspruch 10.
  - 13. Transformierter Organismus nach Anspruch 12 enthaltend einen Vektor nach Anspruch 11.
- 30 14. Verfahren zur Herstellung von Riboflavin, dadurch gekennzeichnet, daß ein Organismus gem. einem der Ansprüche 1 bis 5 eingesetzt wird.

5 15. Verfahren zur Herstellung eines Riboflavin produzierenden ein- oder mehrzelligen Organismus, dadurch gekennzeichnet, daß er so verändert wird, daß dessen Enzymaktivität bezüglich der NAD(P)H-Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.

10

- 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß die Veränderung des Organismus mittels gentechnischer Methoden erfolgt.
- 15 17. Verfahren nach einem der Ansprüche 15 oder 16,

  dadurch gekennzeichnet, daß die Veränderung des Organismus

  durch Austausch des Promotors und/oder Erhöhung der

  Genkopienzahl erzielt wird.
- 20 18. Verfahren nach einem der Ansprüche 15 bis 17, dadurch gekennzeichnet, daß durch die Änderung des endogenen Isocitrat-Dehydrogenase-Gens ein Enzym mit erhöhter Aktivität erzeugt wird.
- 25 19. Verwendung des Organismus nach einem der Ansprüche 1 bis 5 sowie 12 und 13 zur Herstellung von Riboflavin.
  - 20. Verwendung des Isocitrat-Dehydrogenase-Gens nach einem der Ansprüche 6 bis 9 und der Gen-Struktur nach Ansprüch 10 zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 sowie 12 und 13.
    - 21. Verwendung des Vektors nach Anspruch 11 zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 sowie 12 und 13.

30



# 5 Zusammenfassung

Die vorliegende Erfindung betrifft einen ein- oder mehrzelligen Organismus, insbesondere Mikroorganismus, zur biotechnischen Herstellung von Riboflavin wobei dessen Enzymaktivität bezüglich der NAD(P)H-Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.

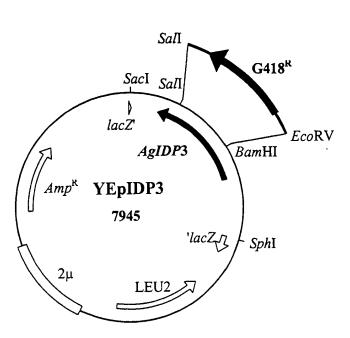


Fig. 1

en in grand de la companya de la co

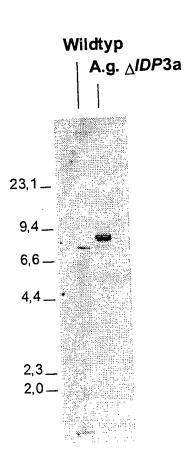


Fig. 2

المراجع وموجو المداري

च रक्षा का नाव्यक्षिक्षकात्रा

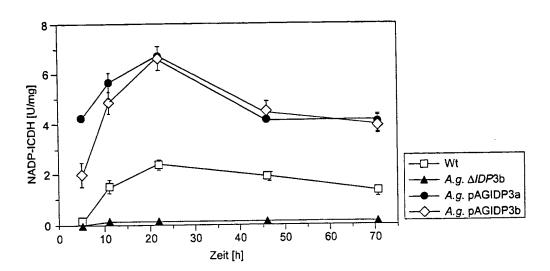
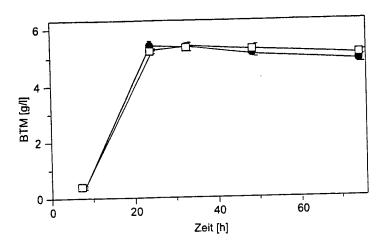


Fig. 3

. . .



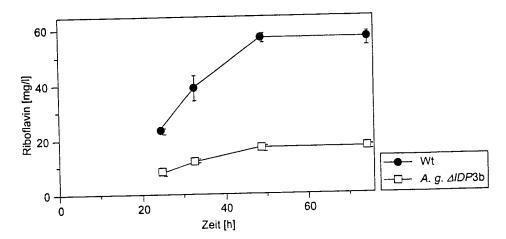


Fig. 4



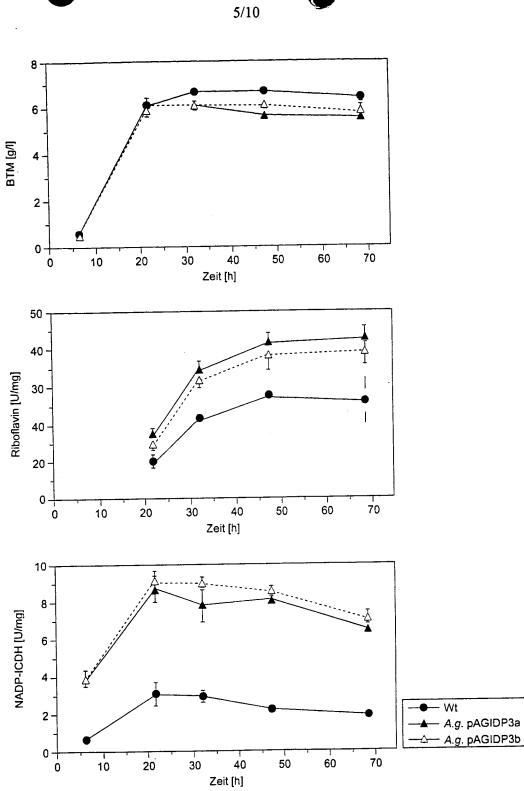
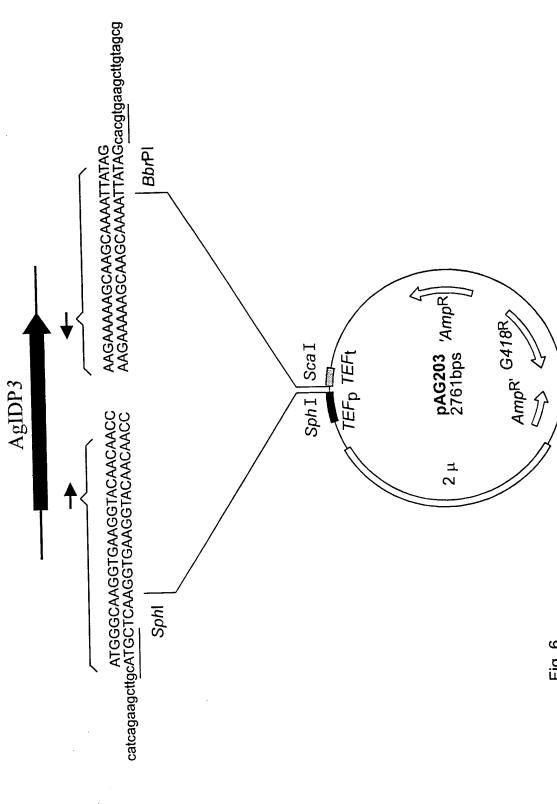
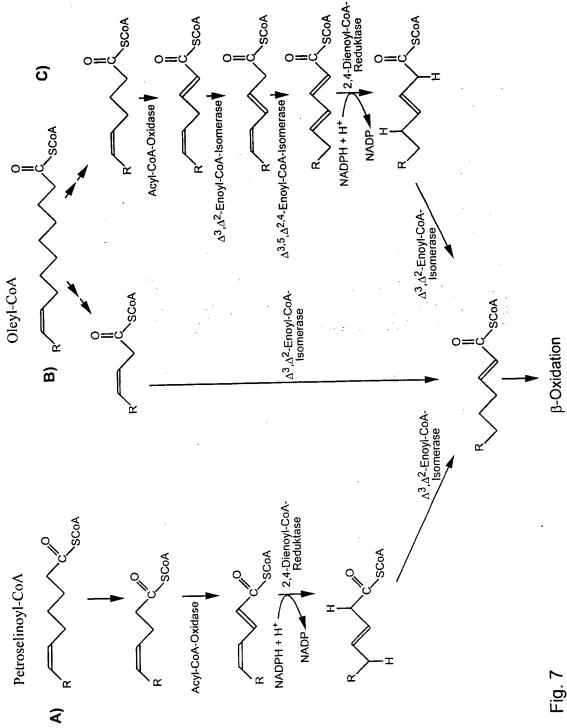


Fig. 5





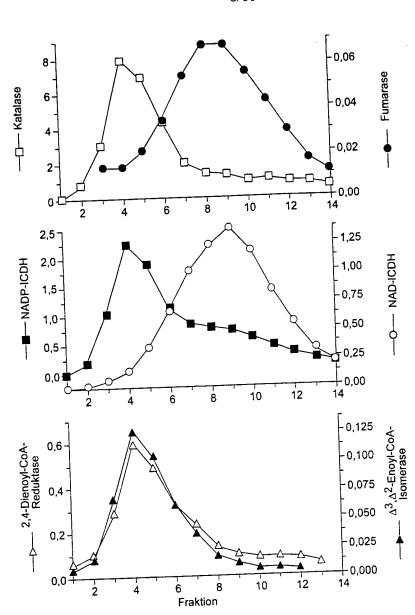
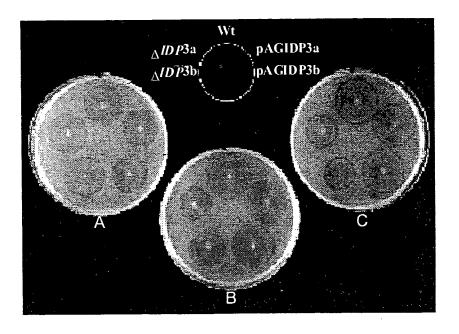


Fig. 8

Secretary of the second



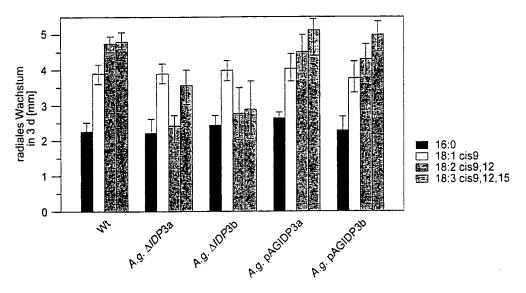


Fig. 9

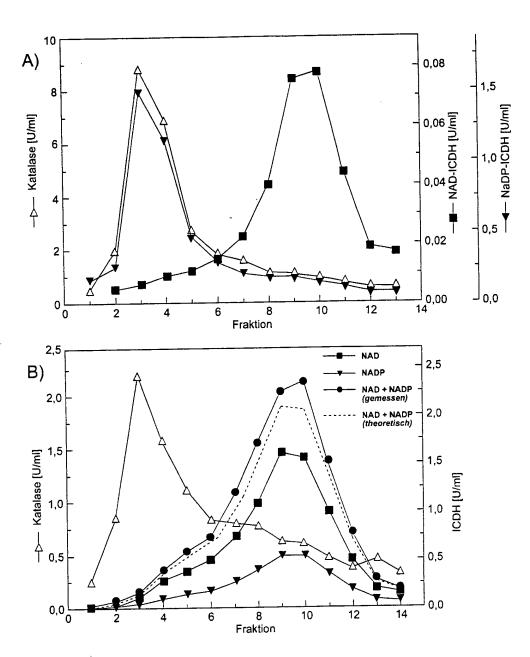


Fig. 10

SEQUENZPROTOKOLL

<110> BASF AG, Forschungszentrum Jülich GmbH

<120> Ein- oder mehrzellige Organismen zur Herstellung von Riboflavin

<130> FZJ9909PCT

<140> PCT/EP00/07370

<141> 2000-07-31

<160> 2

<170> PatentIn Ver. 2.1

<210> 1

<211> 2321

<212> DNA

<213> Pilz der Spezies Ashbya Gossypii

<220>

<221> CDS

<222> (718)..(1266)

ctgcagcaaa tcgaggtgat cgccaacgag gtggacgtgc ggcaggacgg gacctggtgc 60 atccggtacc gcgacgagtc cgagcacggg cacgacaagt cgcggtcgat cgcggcgtgc 120 aagcageget ggcaacacet egageeegeg eeggtgtatt tetaetgegg egatgggate 180 agcgacctga gcgctgcgaa ggaatgcgac ctgctgtttg cgaagagtgg caaggacctg 240 atctccttct gcaagaagca ggacgttccg ttccgcgagt tcaacacttt tgacgatgtg 300 ctgagegegg teaagegegt ggtggeggge gaggeetetg teaeggaact eeagggggge 360 teegetgegt aageaetgte tgeateagtg acettggegg tagetgegat ttgtaaetae 420

ctacgtaatt agtcctgctc gcgctgcggt ccagtgctag gcacgcccca catgaaaggc 480 agccgtaagc aattagtaac ggcctagtac ggctccgatg tatgtgctag cacatgacag 540

cccaacgggt tgagaagtec ggctcgaatc atttccgcgc cgagtgggtc gtgggtggag 600

cogcocgaco cottgtcago gogggcagtt ggatataagg cagtggttgt agcaaaagtg 660

agXtgcgtgca tttcacgaag ccgagcgcaa caacgcacag acatcagtaa gcagct atg 720

ggc aag gtg aag gta caa caa ccc atc gtc gag atg gac ggc gac gaa 768 Gly Lys Val Lys Val Gln Gln Pro Ile Val Glu Met Asp Gly Asp Glu

cag acg cgg atc atc tgg cac ttg atc aag gat cag ctc atc ttc ccc 816 Gln Thr Arg Ile Ile Trp His Leu Ile Lys Asp Gln Leu Ile Phe Pro

and the second of the second of the second

tac ttg gac gtg gac ttg aag tac tac gat ctt tcc att gag aac agg  Tyr Leu Asp Val Asp Leu Lys Tyr Tyr Asp Leu Ser Ile Glu Asn Arg  35  40  45	
gat gcc acc gag gac cgc gtg act gtg gag tct gcg gag gcg acc ctc 912 Asp Ala Thr Glu Asp Arg Val Thr Val Glu Ser Ala Glu Ala Thr Leu 50 55 60 65	
aag tac ggc gtt gcc gtc aag tgt gcg att att acc ccg gac gag gcg 960 Lys Tyr Gly Val Ala Val Lys Cys Ala Ile Ile Thr Pro Asp Glu Ala 70 75 80	
cgt gtc gag gag ttc ggg ctc aag gag atg tgg aag tct ccc aac ggg 1008 Arg Val Glu Glu Phe Gly Leu Lys Glu Met Trp Lys Ser Pro Asn Gly 85 90 95	
acc atc cgg aac atc ctc ggc ggg acc gtc ttc aga gag ccc att att 1056 Thr Ile Arg Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile Ile 100 105 110	
atc cca agg atc ccc aga ctg gtg ccc ggc tgg aac gag ccg atc att 1104  Ile Pro Arg Ile Pro Arg Leu Val Pro Gly Trp Asn Glu Pro Ile Ile  115 120 125	:
gtc ggc aga cac gcg ttt ggg gac cag tac aag gcg acc gac gtt gtc 1152 Val Gly Arg His Ala Phe Gly Asp Gln Tyr Lys Ala Thr Asp Val Val 130 135 140 145	2
att cca ggc gag ggc acg ttg aag ctg gtc ttt gaa agc aag gac ggg 1200 Ile Pro Gly Glu Gly Thr Leu Lys Leu Val Phe Glu Ser Lys Asp Gly 150 155	)
gac aag tcc aag aat ctt gac ctg gag ttc ttt gaa tac ccc aag gat 1248 Asp Lys Ser Lys Asn Leu Asp Leu Glu Phe Phe Glu Tyr Pro Lys Asp 165 170 175	3
ggc ggt gtt gcc atg acc atgtactaca ccaccgactc gatcaccggc 1296 Gly Gly Val Ala Met Thr 180	6
tttgccaagt cgagcttcga gttggcgttg caaagaaaga tgccgctata ttcgacaacg 135	6
aagaacacga tottgaagaa gtacgacggo aagtttaagg atattttcga gggcatgtac 141	
ccageggagt acaaggagaa gtttgagget getggeatet ggtatgaaca eagaetgatt 147	
gacgatatgg ttgcgcagat gttgaagtcc aagggcggct tcatcattgc catgaagaac 153	
tacgatggtg atgtgcagtc ggacatcgtc gcccagggct tcgggtcttt gggtctcatg 159	
acttctgttc ttgtgtctcc agatggaaag accttcgaga gtgaggccgc acatggcact 165	
gtcacccggc actacagaca gcaccagcag ggcaaggaaa catccaccaa ctctattgcc 171	
tetatttttg cetggatgeg eggtattata cacagaggta aggtegaegg taccecagat 177	
gtcgtgaagt tcggcgagtt gttggagaag tccaccctgg acacggtgca ggaggacatc 183	
atgaccaagg acctagcgtt gattttgggc aagaccgaca gagccagcta tgttaccacg 189	<del>)</del> 6

gaagagttta tcacagcagt agcgaaccgc ttagcgatgg ctacaagcgt cttttttgtg 1956 aataagaaaa agcaagcaaa attatagcct aggctgcctg tagcgtctat ttattactag 2016 totagoatat ctagoacaag aatatagata ctgagocato cgcccaggat tacagtcagg 2076 attccaactt gtaaacctcc ggtggtgcgc actcgccgca aattaggtga gcttgccatt 2136 agtcatccga ggcgcagaat gagtagggtt tatagtaaac ccgggtgctg taacaccaga 2196 teccaetttt cetggeacag tatttttgee gacaacggea etgetaaceg tttetcaaet 2256 acgcgcaata atgtaggtcg cacggtccga tgaaaactaa tgcgcagtag catgacatgg 2316 2321 aattc

<210> 2

<211> 183

<212> PRT

<213> Pilz der Spezies Ashbya Gossypii

Met Gly Lys Val Lys Val Gln Gln Pro Ile Val Glu Met Asp Gly Asp

Glu Gln Thr Arg Ile Ile Trp His Leu Ile Lys Asp Gln Leu Ile Phe

Pro Tyr Leu Asp Val Asp Leu Lys Tyr Tyr Asp Leu Ser Ile Glu Asn

Arg Asp Ala Thr Glu Asp Arg Val Thr Val Glu Ser Ala Glu Ala Thr

Leu Lys Tyr Gly Val Ala Val Lys Cys Ala Ile Ile Thr Pro Asp Glu

Ala Arg Val Glu Glu Phe Gly Leu Lys Glu Met Trp Lys Ser Pro Asn

Gly Thr Ile Arg Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile

Ile Ile Pro Arg Ile Pro Arg Leu Val Pro Gly Trp Asn Glu Pro Ile 120

Ile Val Gly Arg His Ala Phe Gly Asp Gln Tyr Lys Ala Thr Asp Val 135 130

Val Ile Pro Gly Glu Gly Thr Leu Lys Leu Val Phe Glu Ser Lys Asp

Gly Asp Lys Ser Lys Asn Leu Asp Leu Glu Phe Phe Glu Tyr Pro Lys 165

Asp Gly Gly Val Ala Met Thr 180

### (12) NACH DEM VERTRE-UBER DIE INTERNATIONALE ZUSAMMENA-ZEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum Internationales Büro





(43) Internationales Veröffentlichungsdatum 15. Februar 2001 (15.02.2001)

**PCT** 

# (10) Internationale Veröffentlichungsnummer WO 01/11052 A2

(51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/53, 15/80, 9/04, 1/15, C12P 25/00 // (C12N 1/15, C12R 1:645)

20, D-52428 Jülich (DE). MAETING, Ines [DE/DE]; Wiesenstr. 3, D-52428 Jülich (DE).

(21) Internationales Aktenzeichen:

PCT/EP00/07370

(22) Internationales Anmeldedatum:

31. Juli 2000 (31.07.2000)

(25) Einreichungssprache:

Deutsch

(26) Veröffentlichungssprache:

Deutsch

(30) Angaben zur Priorität: 199 37 548.8 9. August 1999 (09.08.1999) DB

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; D-67056 Ludwigshafen (DE). FORSCHUNGSZENTRUM JÜLICH GMBH [DE/DE]; D-52425 Jülich (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): ALTHÖFER, Henning [DE/DE]; Mainstrasse 12, D-67117 Limburgerhof (DE). ZELDER, Oskar [DE/DE]; Rossmarktstrasse 27, D-67346 Speyer (DE). REVUELTA DOVAL, Jose L. [ES/ES]; Grillo 11 4E, E-37001 Salamanca (ES). SANTOS GARCIA, Maria Angeles [ES/ES]; C/Escuelas, 1-5, E-37001 Salamanca (ES). SAHM, Hermann [DE/DE]; Wendelinusstrasse 71, D-52428 Jülich (DE). STAHMANN, Klaus-Peter [DE/DE]; Wilhelmstrasse

(74) Anwalt: FITZNER, Uwe; Lintorfer Strasse 10, D-40878 Ratingen (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Veröffentlicht:

 Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

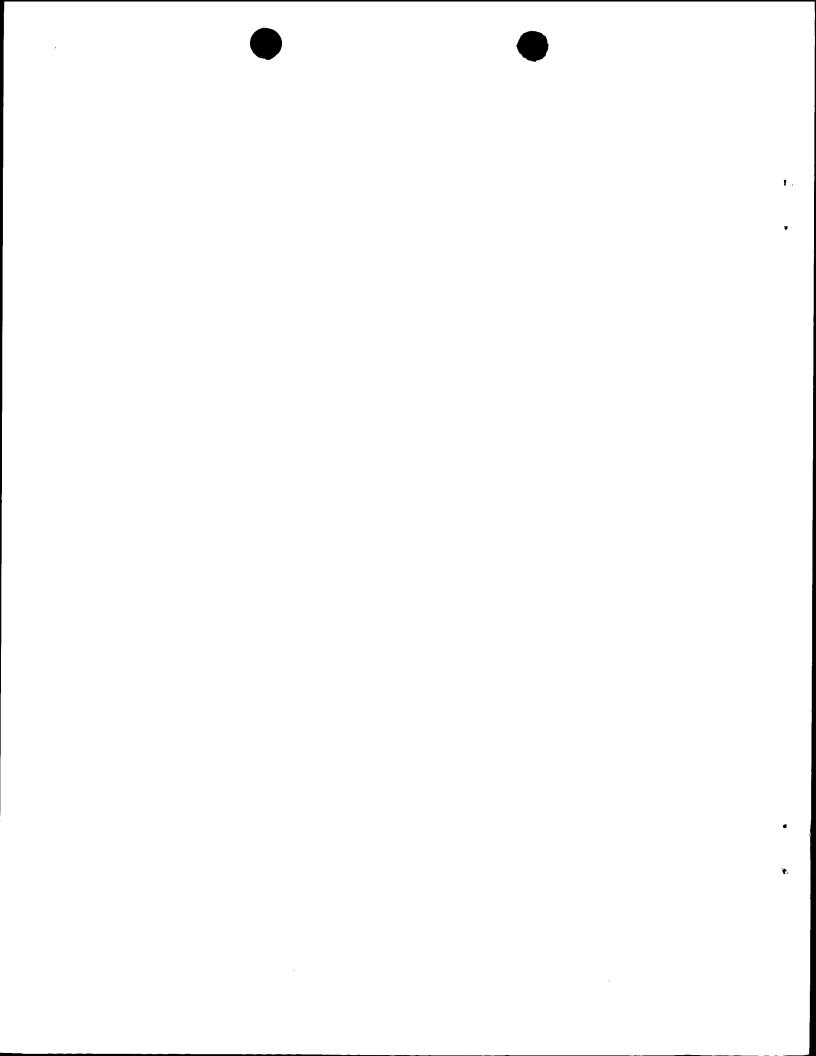
(54) Title: MONOCELLULAR OR MULTICELLULAR ORGANISMS FOR THE PRODUCTION OF RIBOFLAVIN

(54) Bezeichnung: EIN- ODER MEHRZELLIGE ORGANISMEN ZUR HERSTELLUNG VON RIBOFLAVIN

(57) Abstract: The invention relates to a monocellular or multicellular organism, especially a microorganism, for biotechnological production of riboflavin, whereby the enzymatic activity thereof with respect to NAD(P)H formation is higher than that of a wild type of species Ashbya gossypii ATCC10895.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft einen ein- oder mehrzelligen Organismus, insbesondere Mikroorganismus, zur biotechnischen Herstellung von Riboflavin, wobei dessen Enzymaktivität bezüglich der NAD(P)H Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.





## 5 Ein- oder mehrzellige Organismen zur Herstellung von Riboflavin

Die vorliegende Erfindung betrifft einen ein- oder mehrzelligen Organismus zur Herstellung von Riboflavin.

Das Vitamin B2, auch Riboflavin genannt, ist für Mensch und Tier 10 essentiell. Bei Vitamin-B2-Mangel treten Entzündungen der Mund- und Rachenschleimhäute. Risse in den Mundwinkeln. Juckreiz Entzündungen in den Hautfalten u. a. Hautschäden, Bindehautentzündungen, verminderte Sehschärfe und Trübung der Hornhaut auf. Bei Säuglingen und Kindern können Wachstumsstillstand und Gewichts-15 abnahme eintreten. Das Vitamin B2 hat daher wirtschaftliche Bedeutung insbesondere als Vitaminpräparat bei Vitaminmangel sowie als Futtermittelzusatz. Daneben wird es auch als Lebensmittelfarbstoff, beispielsweise in Mayonnaise, Eiscreme, Pudding etc., eingesetzt.

20

25

30

Die Herstellung von Riboflavin erfolgt entweder chemisch oder mikrobiell. Bei den chemischen Herstellungsverfahren wird das Riboflavin in der Regel in mehrstufigen Prozessen als reines Endprodukt gewonnen, wobei allerdings auch relativ kostspielige Ausgangsprodukte – wie beispielsweise D-Ribose – eingesetzt werden müssen.

Eine Alternative zur chemischen Herstellung des Riboflavins bietet die Herstellung dieses Stoffes durch Mikroorganismen. Die mikrobielle Herstellung des Riboflavins eignet sich insbesondere für solche Fälle, in denen eine hohe Reinheit dieser Substanz nicht erforderlich ist. Dies ist beispielsweise dann der Fall, wenn das Riboflavin als Zusatz zu Futtermittelprodukten eingesetzt werden soll. In solchen Fällen hat die mikrobielle Herstellung des Riboflavins den Vorteil, daß diese Substanz in einem einstufigen Prozeß gewinnbar ist. Auch können als

20

25

30

5 Ausgangsprodukte für die mikrobielle Synthese nachwachsende Rohstoffe, wie beispielsweise pflanzliche Öle, eingesetzt werden.

Die Herstellung von Riboflavin durch Fermentation von Pilzen wie Ashbya gossypii oder Eremothecium ashbyi ist bekannt (The Merck Index, Windholz et al., eds. Merck & Co., Seite 1183, 1983, A. Bacher, F. Lingens, Augen. Chem. 1969, S. 393); aber auch Hefen, wie z.B. Candida oder Saccharomyces, und Bakterien wie Clostridium, Bacillus und Corynebakterium sind zur Riboflavinproduktion geeignet.

Zudem sind Verfahren mit der Hefe Candida famata beispielsweise in der US 05231007 beschrieben.

Riboflavin-überproduzierende Bakterienstämme sind beispielsweise in der EP 405370 beschrieben, wobei die Stämme durch Transformation der Riboflavin- Biosynthese-Gene aus Bacillus subtilis erhalten wurden. Diese Prokaryonten-Gene waren aber für ein rekombinantes Riboflavin-Herstellungsverfahren mit Eukaryonten wie Saccharomyces cerevisiae oder Ashbya gossypii ungeeignet. Daher wurden gemäß der WO 93/03183 für die Riboflavin-Biosynthese spezifische Gene aus einem Eukaryonten, nämlich aus Saccharomyces cerevisiae, isoliert, um damit ein rekombinantes Herstellungsverfahren für Riboflavin in einem eukaryontischen Produktionsorganismus bereitzustellen. Derartige rekombinante Herstellungsverfahren haben für die Riboflavin-Produktion jedoch dann keinen oder nur begrenzten Erfolg, wenn die Bereitstellung von Substrat für die an der Riboflavin-Biosynthese spezifisch beteiligten Enzyme unzureichend ist.

1967 fand Hanson (Hanson AM, 1967, in Microbial Technology, Peppler,
 HJ, pp.222-250 New York), daß der Zusatz der Aminosäure Glycin die
 Riboflavin-Bildung von Ashbya gossypii steigert. Ein derartiges Verfahren

ist jedoch nachteilig, weil Glycin ein sehr teurer Rohstoff ist. Aus diesem 5 Grunde war man bestrebt, durch Herstellung von Mutanten die Riboflavin-Produktion zu optimieren.

Aus der DE 19545468.5 A1 ist ein weiteres Verfahren zur mikrobiellen Herstellung von Riboflavin bekannt, bei dem die Isocitratlyase-Aktivität oder die Isocitratlyase-Genexpression eines Riboflavin produzierenden Mikroorganismus erhöht ist. Darüber hinaus ist aus der DE 19840709 A1 ein ein- oder mehrzelliger Organismus insbesondere ein Mikroorganismus zur biotechnischen Herstellung von Riboflavin bekannt. Dieser zeichnet sich dadurch aus, daß einen derart veränderten Glycinstoffwechsel 15 aufweist, daß seine Riboflavinsyntheseleistung ohne externe Zufuhr von Glycin mindestens gleich derjenigen eines Wildtyps der Species Ashbya gossypii ATCC10892 ist.

20 Aber auch im Vergleich zu diesen Verfahren besteht noch ein Bedarf zu einer weiteren Optimierung der Riboflavin-Herstellung.

Aufgabe der vorliegenden Erfindung ist es demgemäß, einen ein- oder mehrzelligen Organismus, vorzugsweise einen Mikroorganismus, für die biotechnische Herstellung von Riboflavin zur Verfügung zu stellen, der eine weitere Optimierung der Riboflavin-Bildung ermöglicht. Insbesondere sollte ein Organismus zur Verfügung gestellt werden, der eine Produktion ermöglicht, die gegenüber dem bisherigen Stand der Technik wirtschaftlicher ist. Vor allem soll der Organismus eine im Vergleich zu den bisherigen Organismen erhöhte Riboflavin-Bildung erlauben.

Diese Aufgabe wird durch einen ein- oder mehrzelligen Organismus gelöst, dessen Enzymaktivität der bezüglich NAD(P)H-Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.

25

30

25

30

35

Das Ziel einer beschleunigten intrazellulären NAD(P)H-Versorgung kann durch Erhöhung der Aktivität eines NAD(P)H-bildenden bzw. Senkung der Aktivität eines NAD(P)H verbrauchenden Enzyms bzw. eine Änderung der Spezifität erreicht werden. Dies läßt sich mit den bekannten Methoden der Stammverbesserung von Organismen erreicht werden. D. h. im einfachsten Falle lassen sich entsprechende Stämme nach der in der Mikrobiologie üblichen Selektion mittels Screening herstellen. Ebenso ist die Mutation mit anschließender Selektion einsetzbar. Die Mutation kann hierbei sowohl mittels chemischer als auch mittels physikalischer Mutagenese ausgeführt werden. Eine weitere Methode ist die Selektion und Mutation mit anschließender Rekombination. Schließlich lassen sich die erfindungsgemäßen Organismen mittels Genmanipulation herstellen.

Erfindungsgemäß wird der Organismus derart verändert, daß er intrazellulär NAD(P)H in einer Menge erzeugt, die größer als sein Bedarf für die Aufrechterhaltung seines Metabolismus ist. Diese Erhöhung der intrazellulären NAD(P)H-Erzeugung läßt sich erfindungsgemäß vorzugsweise dadurch erreichen, daß ein Organismus hergestellt wird, bei dem die Enzymaktivität der Isocitrat-Dehydrogenase erhöht ist. Dies kann beispielsweise dadurch erreicht werden, daß durch Veränderung des katalytischen Zentrums ein erhöhter Substratumsatz erfolgt oder indem die Wirkung von Enzyminhibitoren aufgehoben wird. Auch kann eine erhöhte Enzymaktivität der Isocitrat-Dehydrogenase durch Erhöhung der Enzymsynthese, beispielsweise durch Genamplifikation oder durch Ausschaltung von Faktoren, die die Enzym-Biosynthese reprimieren, hervorgerufen werden.

Die Isocitrat-Dehydrogenase-Aktivität kann erfindungsgemäß vorzugsweise durch Mutation des Isocitrat-Dehydrogenase-Gens erhöht werden. Derartige Mutationen können entweder nach klassischen Methoden ungerichtet erzeugt werden, wie beispielsweise durch UV-

5 Bestrahlung oder mutationsauslösende Chemikalien, oder gezielt mittels gentechnologischer Methoden, wie Deletion, Insertion und/oder Nukleotid-Austausch.

Die Isocitrat-Dehydrogenase-Genexpression kann durch Einbau von Isocitrat-Dehydrogenase-Genkopien und/oder durch Verstärkung regulatorischer Faktoren, die die Isocitrat-Dehydrogenase-Genexpression positiv beeinflussen, erreicht werden. So kann eine Verstärkung regulatorischer Elemente vorzugsweise auf Transcriptionsebene erfolgen, indem insbesondere die Transcriptionssignale erhöht werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der m-RNA verbessert wird.

Zur Erhöhung der Genkopienzahl kann beispielsweise das Isocitrat-Dehydrogenase-Gen in ein Genkonstrukt bzw. in einen Vektor eingebaut werden, der vorzugsweise dem Isocitrat-Dehydrogenase-Gen zugeordnete regulatorische Gensequenzen enthält, insbesondere solche, die die Genexpression verstärken. Anschließend wird ein Riboflavin-produzierender Mikroorganismus, mit dem das Isocitrat-Dehydrogenase-Gen enthaltenden Genkonstrukt transformiert.

25

30

20

Erfindungsgemäß kann die Überexpression der Isocitrat-Dehydrogenase auch durch Austausch des Promotors erzielt werden. Hierbei ist es möglich, die höhere enzymatische Aktivität alternativ durch Einbau von Genkopien oder durch Austausch des Promotors zu erzielen. Gleichermaßen ist es jedoch auch möglich, durch gleichzeitigen Austausch des Promotors und Einbau von Genkopien die gewünschte Änderung der Enzymaktivität zu erzielen.

Die Veränderung des Isocitrat-Dehydrogenase-Gens führt zu einer 35 beschleunigten NAD(P)H-Bildung und zugleich zu einer überraschend

5 hohen Steigerung der Riboflavin-Bildung, wie sie bisher nicht erreichbar war.

Das Isocitrat-Dehydrogenase-Gen wird vorzugsweise aus Mikroorganismen, besonders bevorzugt aus Pilzen, isoliert. Dabei sind Pilze der Gattung Ashbya wiederum bevorzugt. Höchst bevorzugt ist die Spezies Ashbya gossypii.

Für die Isolierung des Gens kommen aber auch alle weiteren Organismen, deren Zellen die Sequenz zur Bildung der Isocitrat-Dehydrogenase enthalten, also auch pflanzliche und tierische Zellen, in Betracht. Die 15 Isolierung kann des Gens durch homologe oder Komplementation einer im Isocitrat-Dehydrogenase-Gen defekten Mutante oder auch durch heterologes Probing oder PCR mit heterologen Primern erfolgen. Zur Subklonierung kann das Insert des komplementierenden 20 Plasmids anschließend durch geeignete Schritte mit Restriktionsenzymen in der Größe minimiert werden. Nach Sequenzierung und Identifizierung des putativen Gens erfolgt eine paßgenaue Subklonierung durch PCR. Plasmide, die die so erhaltenen Fragmente als Insert tragen, werden in die Isocitrat-Dehydrogenase-Gen-Defekte Mutante eingebracht, 25 Funktionalität des Isocitrat-Dehydrogenase-Gens getestet wird. Funktionelle Konstrukte werden schließlich zur Transformation eines Riboflavin-Produzenten eingesetzt.

Nach Isolierung und Sequenzierung sind die Isocitrat-Dehydrogenase30 Gene mit Nukleotidsequenzen erhältlich, die für die angegebene Aminosäure-Sequenz oder deren Allelvariation kodieren. Allelvariationen umfassen insbesondere Derivate, die durch Deletion, Insertion und Substitution von Nuldeotiden aus entsprechenden Sequenzen erhältlich sind, wobei die Isocitrat-Dehydrogenase-Aktivität erhalten bleibt. Eine

15

20

25

30

35

7

5 entsprechende Sequenz ist in Figur 2b von Nukleotid 1 bis 1262 angegeben.

Den Isocitrat-Dehydrogenase-Genen ist insbesondere ein Promotor der Nukleotidsequenz von Nukleotid –661 bis –1 gem. Fig. 11 oder eine im wesentlichen gleich wirkende DNA-Sequenz vorgeschaltet. So kann beispielsweise dem Gen ein Promotor vorgeschaltet sein, der sich von dem Promotor mit der angegebenen Nukleotidsequenz durch ein oder mehrere Nukleotidaustausche, durch Insertion und/oder Deletion unterscheidet, ohne daß aber die Funktionalität bzw. die Wirksamkeit des Promotors beeinträchtigt wird. Des weiteren kann der Promotor durch Veränderung seiner Sequenz in seiner Wirksamkeit erhöht oder komplett durch wirksame Promotoren ausgetauscht werden.

Dem Isocitrat-Dehydrogenase-Gen können des weiteren regulatorische Gen-Sequenzen bzw. Regulatorgene zugeordnet sein, die insbesondere die Isocitrat-Dehydrogenase-Gen-Aktivität erhöhen. So können dem Isocitrat-Dehydrogenase-Gen beispielsweise sog. "enhancer" zugeordnet sein, die über eine verbesserte Wechselwirkung zwischen RNA-Polymerase und DNA eine erhöhte Isocitrat-Dehydrogenase-Expression bewirken.

Dem Isocitrat-Dehydrogenase-Gen mit oder ohne vorgeschaltetem Promotor bzw. mit oder ohne Regulator-Gen können ein oder mehrere DNA-Sequenzen vor- und/oder nachgeschaltet sein, so daß das Gen in einer Gen-Struktur enthalten ist. Durch Klonierung des Isocitrat-Dehydrogenase-Gens sind Plasmide bzw. Vektoren erhältlich, die das Isocitrat-Dehydrogenase-Gen enthalten und zur Transformation eines Riboflavin-Produzenten geeignet sind. Die durch Transformation erhältlichen Zellen enthalten das Gen in replizierbarer Form, d.h. in zusätzlichen Kopien auf dem Chromosom, wobei die Genkopien durch

5 homologe Rekombination an beliebigen Stellen des Genoms integriert werden und/oder auf einem Plasmid bzw. Vektor.

Bei den erfindungsgemäß erhaltenen ein- oder mehrzelligen Organismen kann es sich um beliebige für biotechnische Verfahren einsetzbare Zellen handeln. Hierzu zählen beispielsweise Pilze, Hefen, Bakterien sowie pflanzliche und tierische Zellen. Erfindungsgemäß handelt es sich vorzugsweise um transformierte Zellen von Pilzen, besonders bevorzugt von Pilzen der Gattung Ashbya. Hierbei ist die Spezies Ashbya gossypii besonders bevorzugt.

15

35

10

Im folgenden wird die Erfindung näher anhand von Beispielen erläutert, ohne daß damit eine Begrenzung auf den Gegenstand der Beispiele verbunden sein soll:

Das Gen der Isocitrat-Dehydrogenase (IDP3) wurde durch PCR kloniert und dann sequenziert (Sequenz siehe Fig. 11). Die gentechnisch durchgeführte partielle Deletion des Gens durch Austauschmutagenese mit einem Geneticinresistenz-Gen (Fig. 1) wurde durch Southem Blot (Fig. 2) bestätigt. Diese Disruption, d.h. Zerstörung des Gens im Genom des Pilzes, führt dazu, daß der Pilz die davon kodierte Isocitrat-Dehydrogenase nicht mehr bilden kann. Fig. 3 zeigt die Abnahme der Enzymaktivität im Disruptionsstamm AgΔDP3b im Vergleich zum Wildtyp ATCC 10895. In Präparationen der Peroxisomen konnte gezeigt werden, daß dieses Enzym in diesen Organellen lokalisiert ist (Fig. 10). Während die Enzymaktivität in Wildtyp-Peroxisomen deutlich messbar ist, findet sich in den Peroxisomen des Disruptionsstamms keine Aktivität mehr.

Die Disruption des Gens führt zu einer deutlichen Verminderung der Vitaminbildung im Vergleich zum Elternstamm (Fig. 4). Wird das Gen dagegen unter Steuerung des starken TEF-Promotors auf einem Plasmid

- 5 (Fig. 6) in zusätzlicher Kopie in die Ashbya-Zellen gebracht, ist ein deutlicher Anstieg in der Enzymaktivität und der Riboflavinbildung meßbar (Fig. 5).
- Fig. 7 zeigt, daß bei der Verstoffwechselung ungesättigter Fettsäuren NADPH zwei von drei alternativen Reaktionswegen 10 bei Reduktionsmittel benötigt wird. Die darin involvierte 2,4-Dienoyl-CoA-Reduktase konnte in Zellen von Ashbya ebenfalls in Peroxisomen lokalisiert werden (Fig. 8). Die Disruption des IDP3-Gens sollte nun zu einem verringerten Wachstum der Zellen auf Linolsäure oder Linolensäure führen. Das konnte auch gemessen werden (Fig. 9). Damit zeigt sich, daß 15 die Bedeutung der IDP3 für den Stoffwechsel der Zelle in der NADPH-Bildung liegt.

25

30

## 5 Beschreibung der Figuren

- Fig. 1: Schema der Konstruktion des Vektors pIDPkan für den Genaustausch des chromosomalen *AgIDP*3-Gens gegen eine durch Deletion und Insertion des G418<sup>R</sup>-Gens inaktive Genkopie.
- 10 Fig. 2: Überprüfung der partiellen Deletion und gleichzeitigen Insertion der Geneticin-Resistenz-Kassette am *AgIDP*-Lokus mittels Southern-Blot-Analyse. Genomische, *Sph*I-gespaltene DNA wurde mit einer Digoxygenin-markierten Sonde hybridisiert.
- 15 Fig. 3: Vergleich der Enzymaktivitäten der NADP-spezifischen ICDH vom Ashbya-Wildtyp, der Mutante A.g. ΔIDP3b und den AgIDP-Überexprimierern A.g. pAGIDP3a und A.g. pAGIDP3b bei Wachstum auf Glucose-Vollmedium.
- Fig. 4: Vergleich des Wachstums und der Riboflavinbildung vom 20 Ashbya-Wildtyp und der Mutante A.g. ΔIDP3b bei Wachstum auf Sojaöl-Vollmedium.
  - Fig. 5: Vergleich des Wachstums, der Riboflavinbildung und der NADP-spezifischen ICDH vom *Ashbya*-Wildtyp und den *AgIDP*3-Überexprimierern *A.g.* pAGIDP3a und *A.g.* pAGIDP3b bei Kultivierung auf Sojaöl-Vollmedium.
  - Fig. 6: Plasmid zur Überexpression des *AgIDP*3-Gens unter Kontrolle von *TEF*-Promotor und *TEF*-Terminator.

    Zur Einführung der *Sph*I-Schnittstelle war eine Änderung der für die zweite Aminosäure kodierenden Nukleotidsequenz notwendig. Es wurde ein konservativer Austausch der Aminosäure Glycin in Leucin vorgenommen.

5

15

- Fig. 7: Abbauwege ungesättigter Fettsäuren mit Doppelbindungen an geraden (A) und ungeraden (B, C) C-Atomen in Peroxisomen nach Henke et al. (1998).
- 10 Fig. 8 Trennung von aus *Ashbya*-Wildtyp isolierten Organellen im Percoll-Dichtegradienten:

Aktivitäten [U/ml] der Markerenzyme Katalase (Peroxisomen) und Fumarase (Mitochondrien), der NAD- und der NADP- spezifischen ICDH und der für den Abbau ungesättigter Fettsäuren notwendigen 2,4-Dienoyl-CoA-Reduktase und  $\Delta^3,\Delta^2$ - Enoyl-CoA-Isomerase

- Fig. 9: Vergleich des radialen Wachstums von Asbya-Wildtyp, der Mutanten A.g. ΔIDP3a und A.g. ΔIDP3b und den
  20 Überexprimierern A.g. pAGIDP3a und A.g. pAGIDP3b auf verschiedene Fettsäuren (A: 18:1 cis9; b: 18:2 cis9,12; C: 18:3 cis9,12,15).
- Fig. 10: Verteilung der Enzyme Katalase und ICDH im Percoll-Dichtegradienten nach Zentrifugation von Organellen aus Myzel des Wildtyps (A) und der Mutante *A.g.* Δ*IDP*3b (B).

### Beschreibung der Sequenz

30

Nukleotidsequenz und daraus abgeleitete Aminosäuresequenz des für die peroxisomale NADP-spezifische Isocitrat-Dehydrogenase codierenden *AgIDP*3-Gens aus *A. gossypii*.

# 5 Ansprüche

- Ein- oder mehrzelliger Organismus, insbesondere Mikroorganismus, zur biotechnischen Herstellung von Riboflavin, dadurch gekennzeichnet, daß dessen Enzymaktivität bezüglich der NAD(P)H-Bildung höher ist als diejenige eines Wildtyps der Species Ashbya gossypii ATCC10895.
- Ein- oder mehrzelliger Organismus nach Anspruch 1,
   dadurch gekennzeichnet, daß er eine erhöhte Isocitrat Dehydrogenase-Aktivität aufweist.
  - Ein- oder mehrzelliger Organismus nach einem der Ansprüche 1 oder 2,
     dadurch gekennzeichnet, daß er ein Pilz ist.

20

 Ein- oder mehrzelliger Organismus nach einem der Ansprüche 1 bis 3,
 dadurch gekennzeichnet, daß er ein Pilz aus der Gattung Ashbya ist.

25

 Ein- oder mehrzelliger Organismus nach einem der Ansprüche 1 bis 4,
 dadurch gekennzeichnet, daß er ein Pilz der Spezies Ashbya gossypii ist.

30

6. Isotitrat-Dehydrogenase-Gen mit einer für die in Fig. 11 angegebenen Aminosäuresequenz und deren Allelvariation kodierenden Nukleotidsequenz.

- 5 7. Isocitrat-Dehydrogenase-Gen nach Anspruch 6 mit der Nukleotidsequenz von Nukleotid 1 bis 1262 gem. der Fig. 11 oder einer im wesentlichen gleich wirkenden DNA-Sequenz.
- Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 oder 7
   mit einem vorgeschalteten Promotor der Nukleotidsequenz mit Nukleotid –661 bis –1 gem. der Fig. 11 oder einer im wesentlichen gleich wirkenden DNA-Sequenz.
- lsocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 8 mit
   diesem zugeordneten regulatorischen Gensequenzen.
  - 10. Gen-Struktur enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 9.
- 20 11. Vektor enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 9 oder eine Gen-Struktur nach Anspruch 10.
- 12. Transformierter Organismus zur Herstellung von Riboflavin enthaltend in replizierbarer Form ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 9 oder eine Gen-Struktur nach Anspruch 10.
  - 13. Transformierter Organismus nach Anspruch 12 enthaltend einen Vektor nach Anspruch 11.
- 30 14. Verfahren zur Herstellung von Riboflavin, dadurch gekennzeichnet, daß ein Organismus gem. einem der Ansprüche 1 bis 5 eingesetzt wird.

Verfahren zur Herstellung eines Riboflavin produzierenden ein- oder mehrzelligen Organismus, dadurch gekennzeichnet, daß er so verändert wird, daß dessen Enzymaktivität bezüglich der NAD(P)H-Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.

10

- 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß die Veränderung des Organismus mittels gentechnischer Methoden erfolgt.
- 15 17. Verfahren nach einem der Ansprüche 15 oder 16,

  dadurch gekennzeichnet, daß die Veränderung des Organismus
  durch Austausch des Promotors und/oder Erhöhung der
  Genkopienzahl erzielt wird.
- 20 18. Verfahren nach einem der Ansprüche 15 bis 17, dadurch gekennzeichnet, daß durch die Änderung des endogenen Isocitrat-Dehydrogenase-Gens ein Enzym mit erhöhter Aktivität erzeugt wird.
- 25 19. Verwendung des Organismus nach einem der Ansprüche 1 bis 5 sowie 12 und 13 zur Herstellung von Riboflavin.
- Verwendung des Isocitrat-Dehydrogenase-Gens nach einem der Ansprüche 6 bis 9 und der Gen-Struktur nach Ansprüch 10 zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 sowie 12 und 13.
  - 21. Verwendung des Vektors nach Anspruch 11 zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 sowie 12 und 13.

1/10

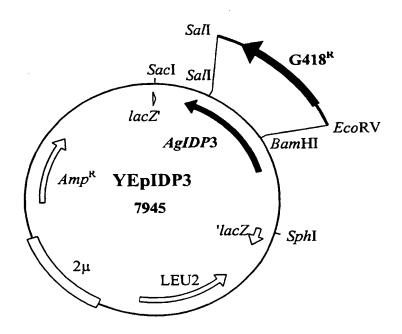
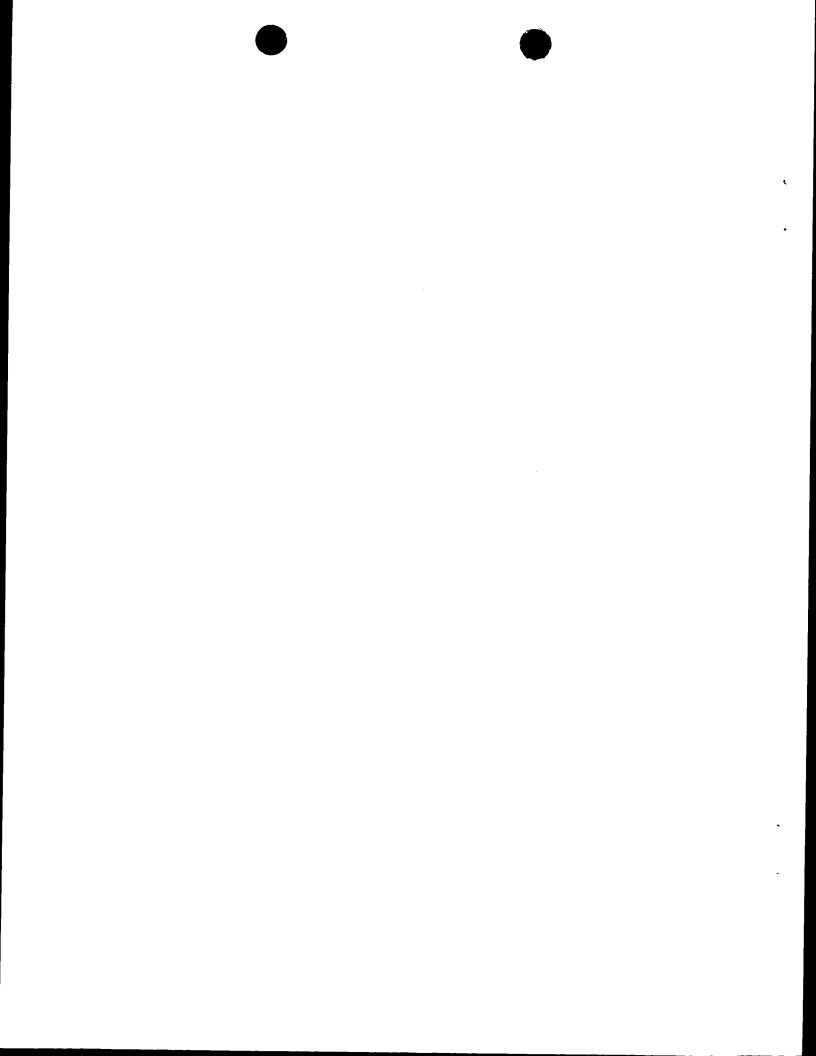


Fig. 1



2/10

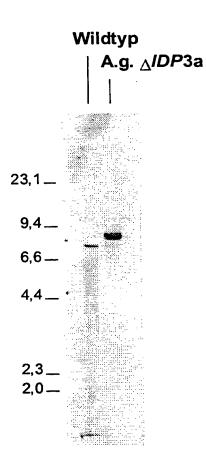
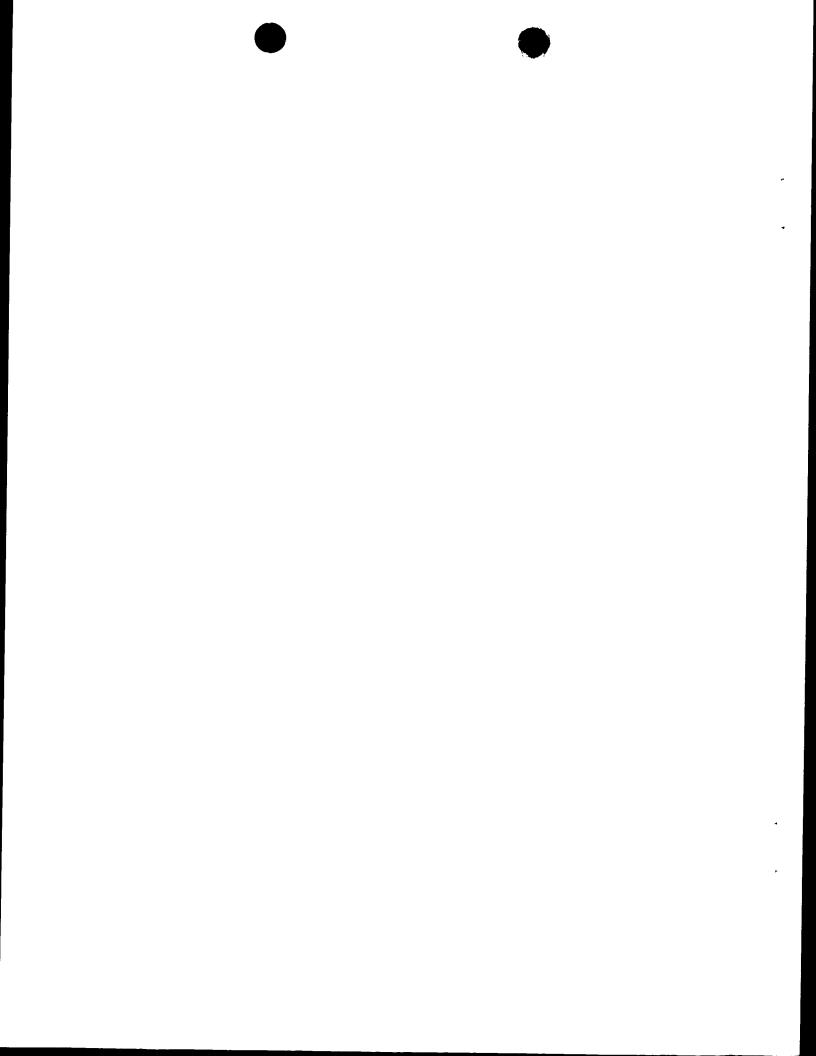


Fig. 2



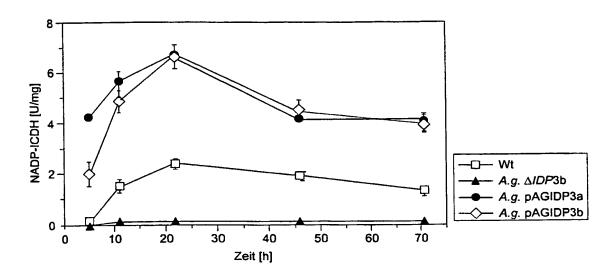
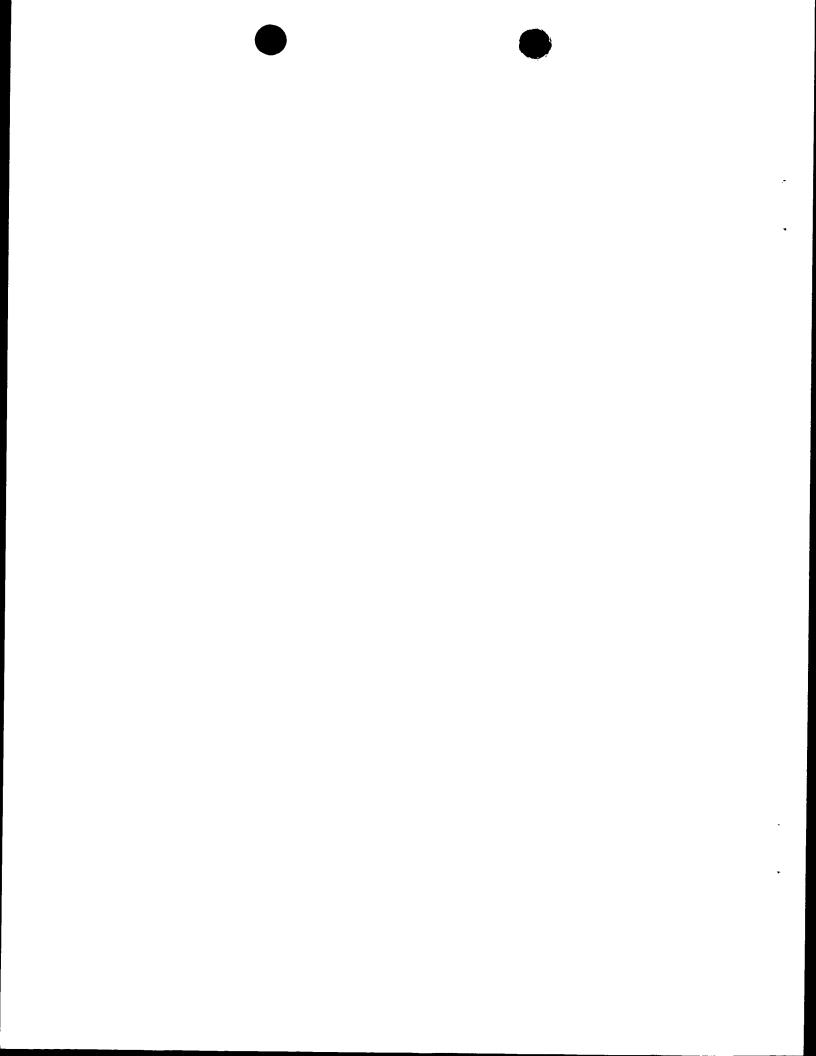
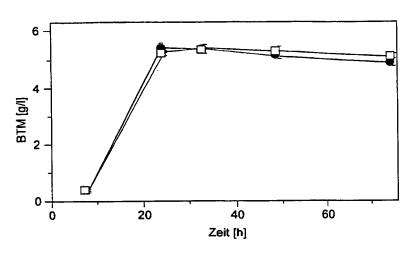


Fig. 3





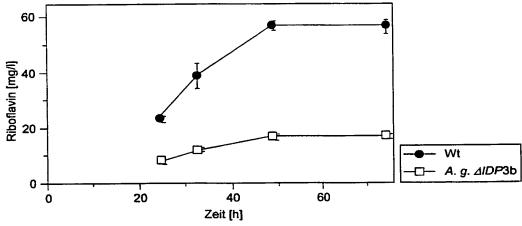
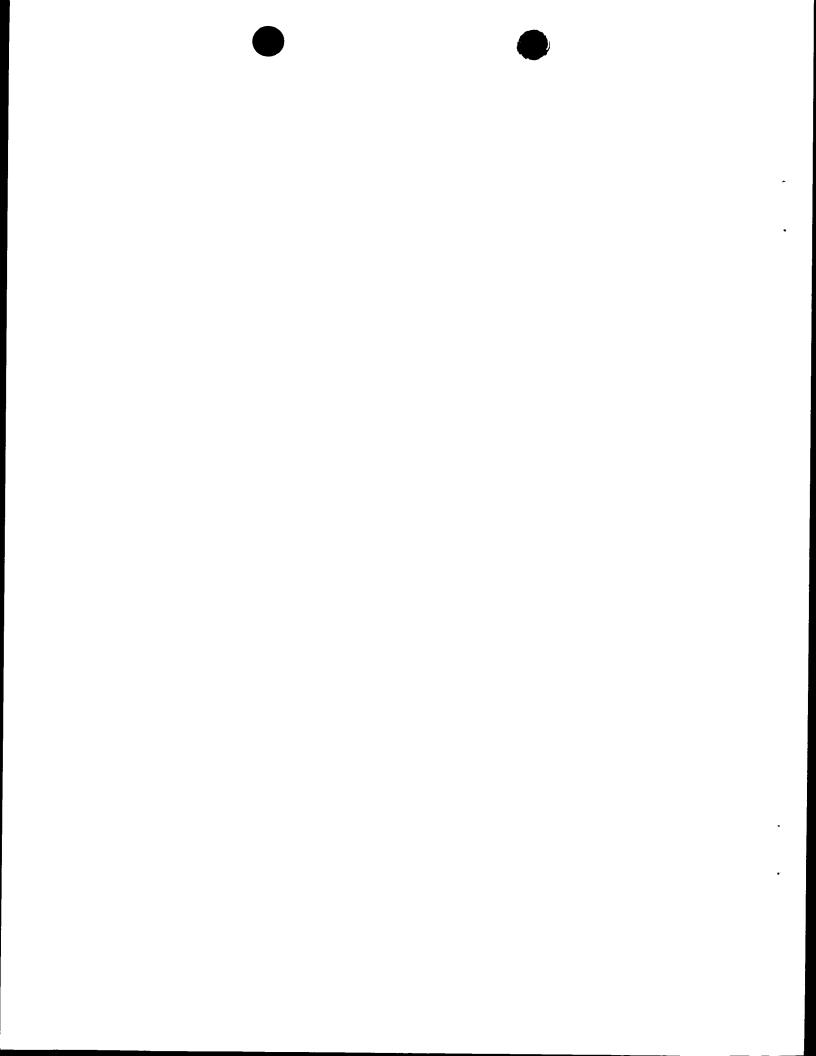


Fig. 4



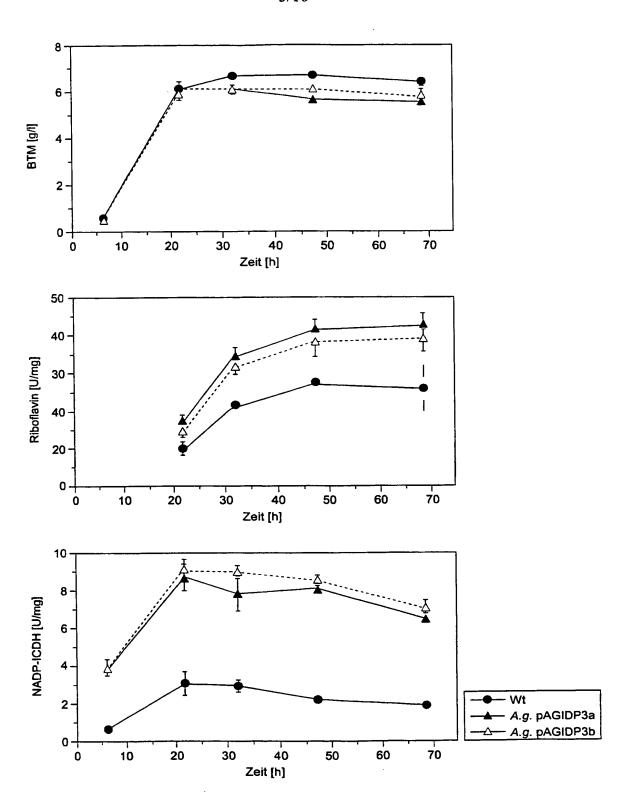
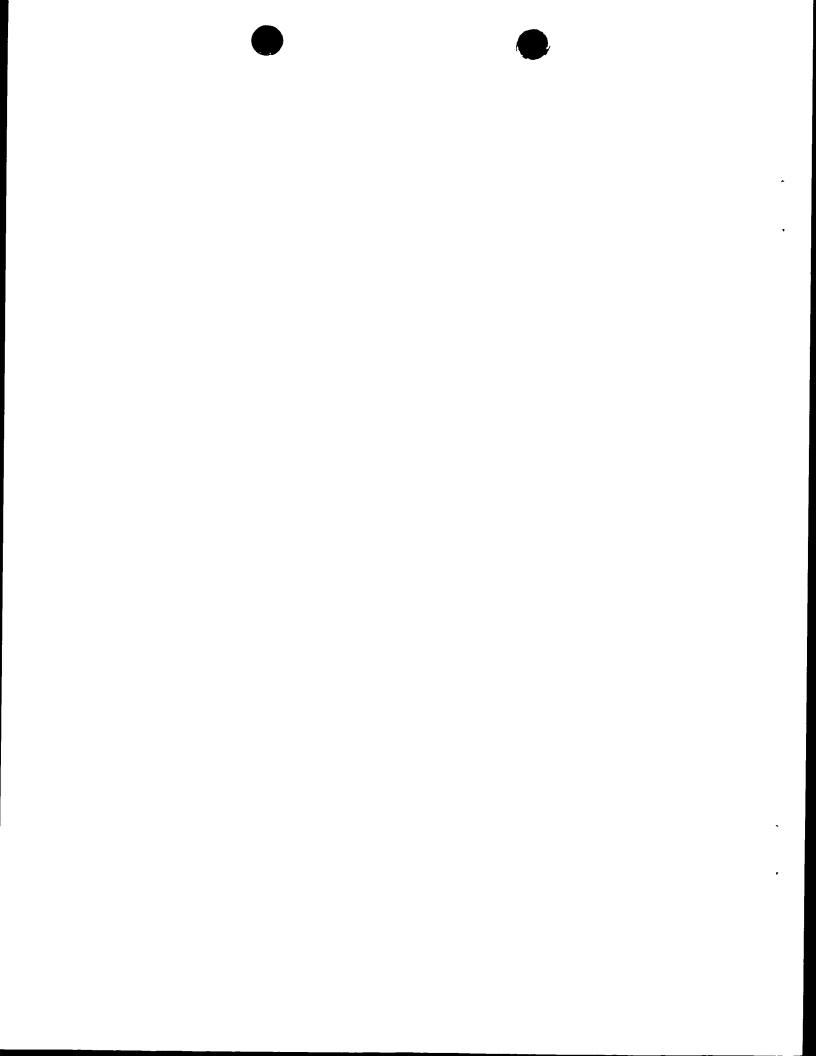
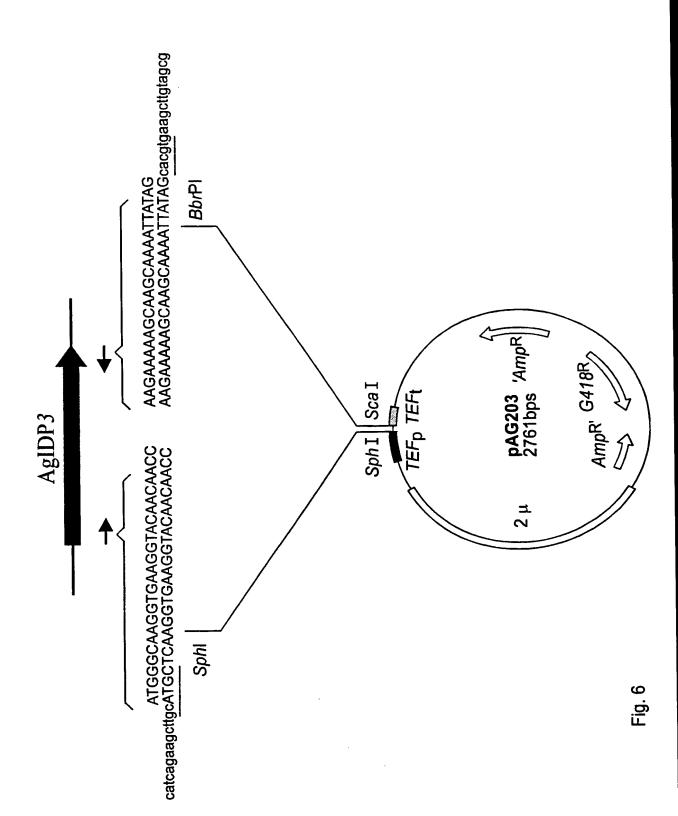
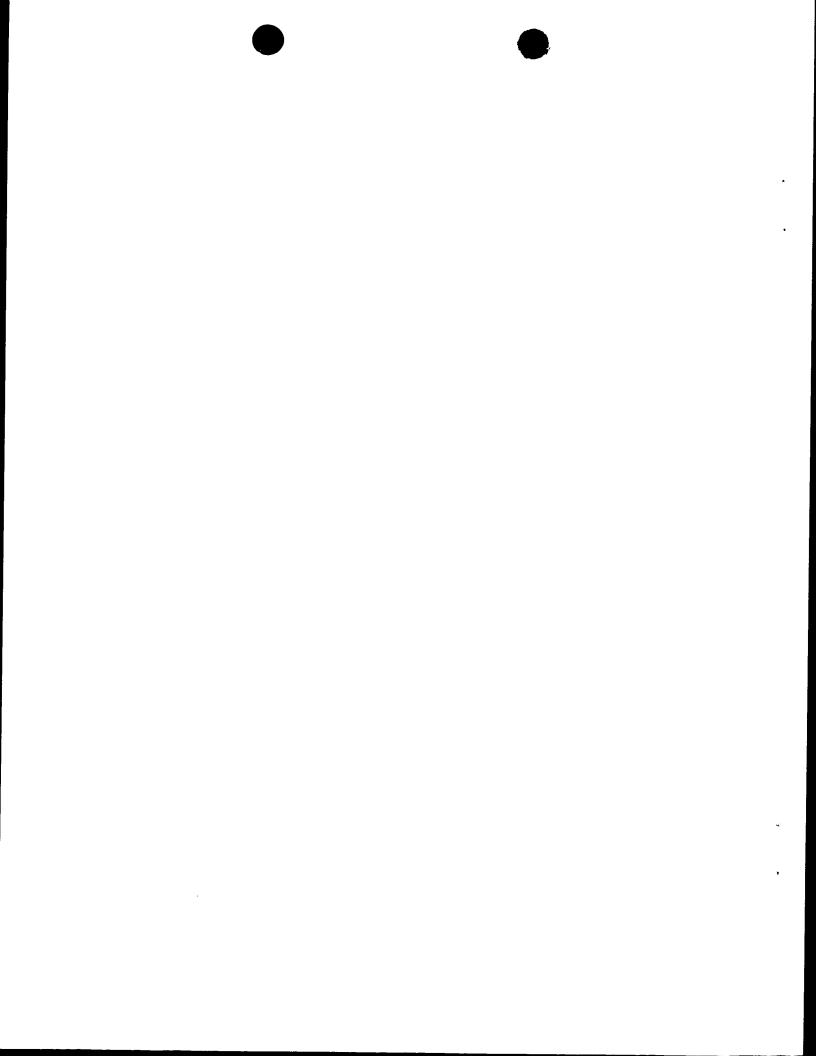
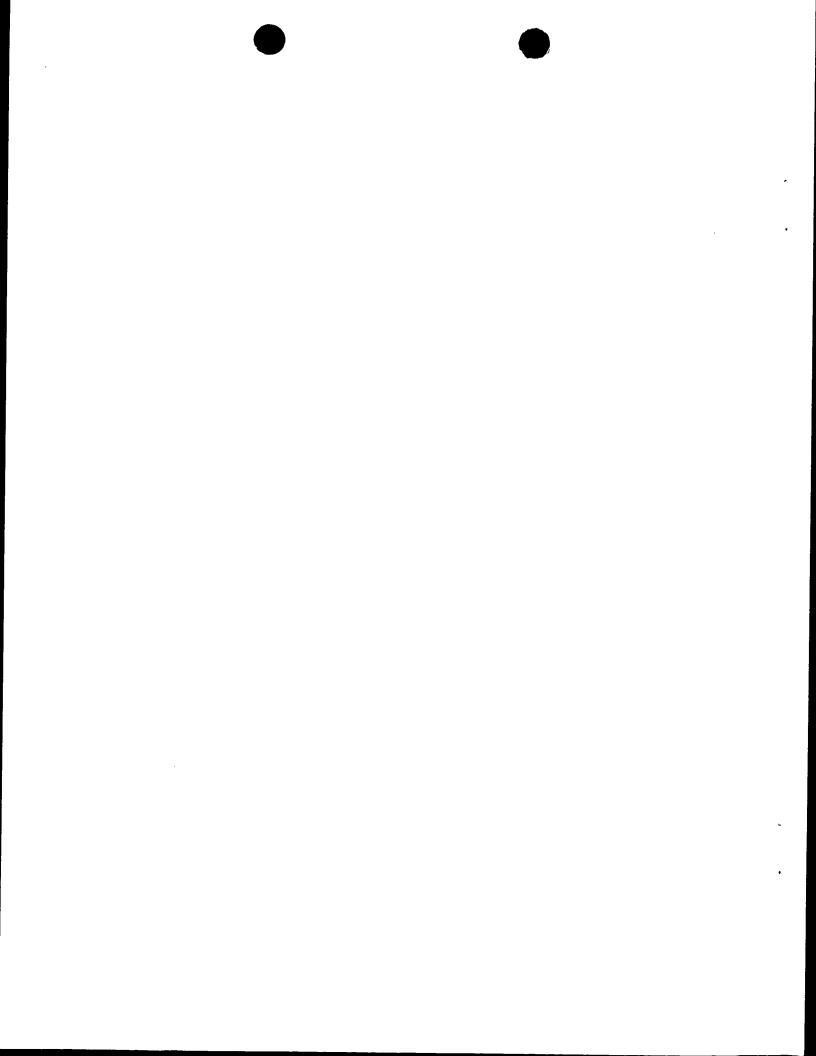


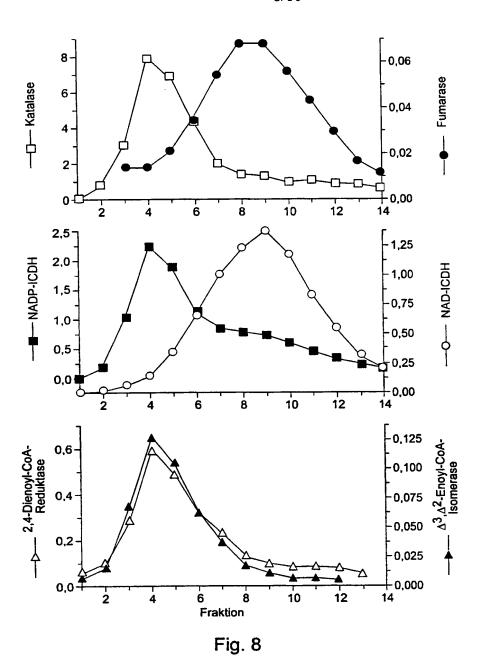
Fig. 5



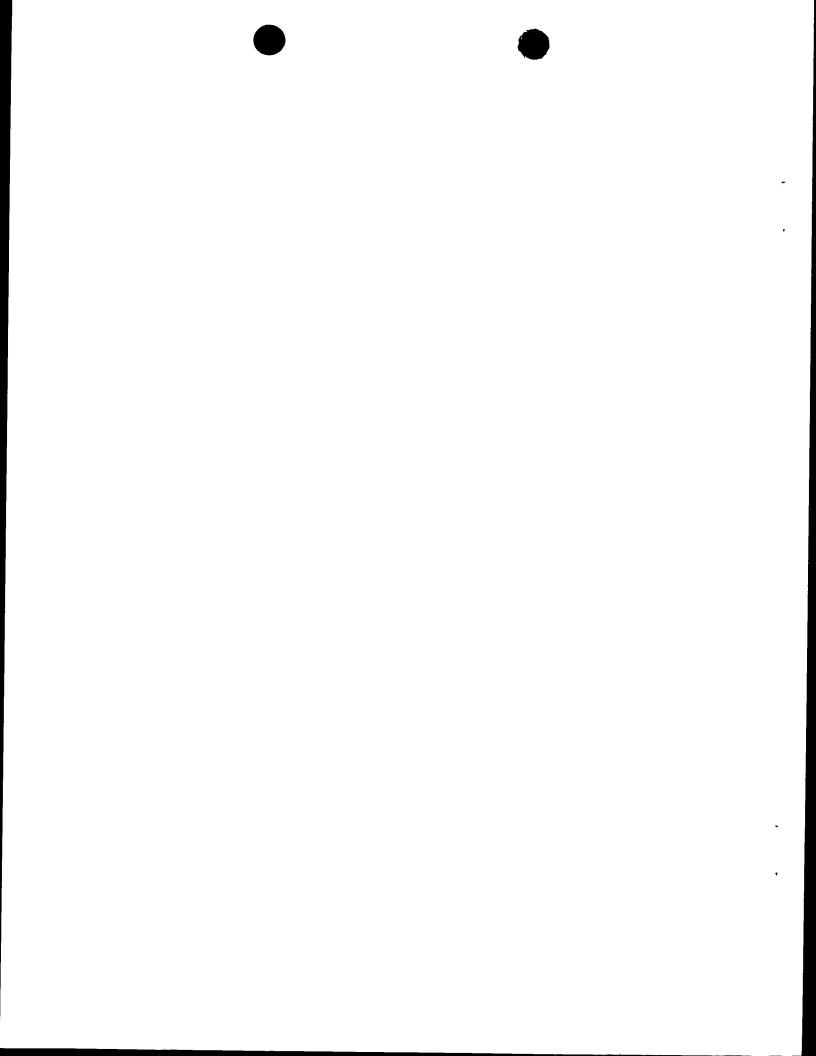


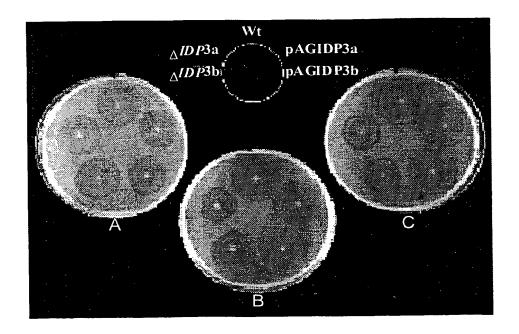






ERSATZBLATT (REGEL 26)





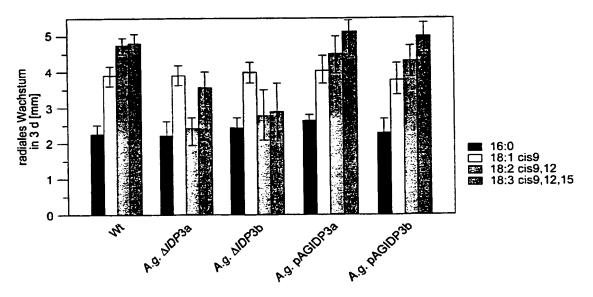
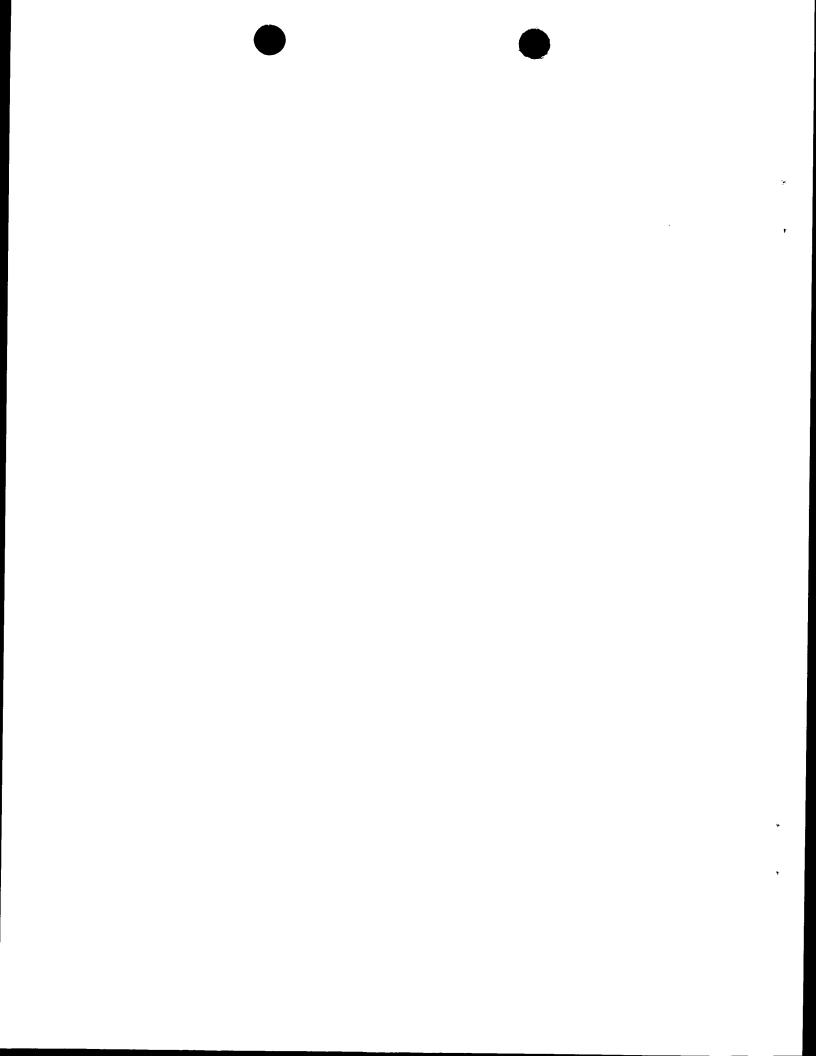


Fig. 9



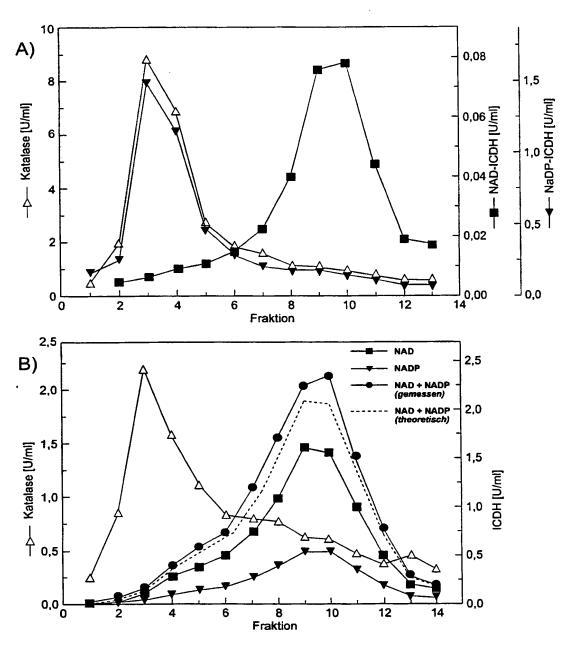
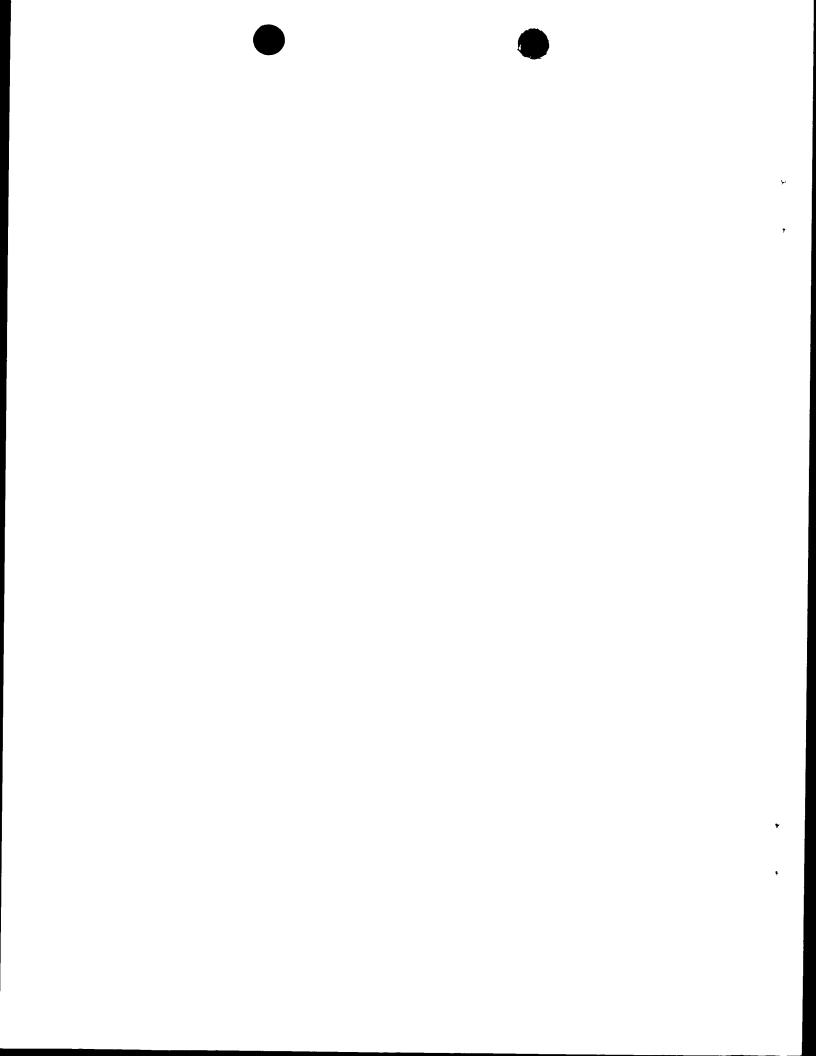


Fig. 10



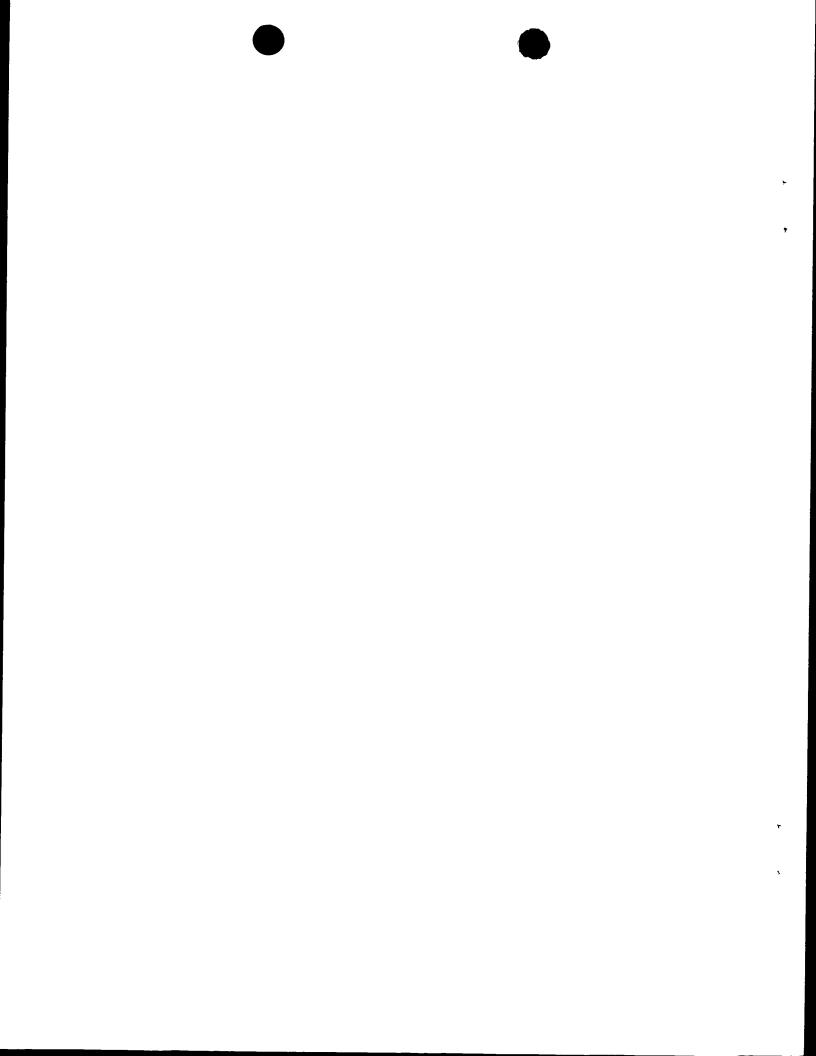
1

#### SEQUENZPROTOKOLL

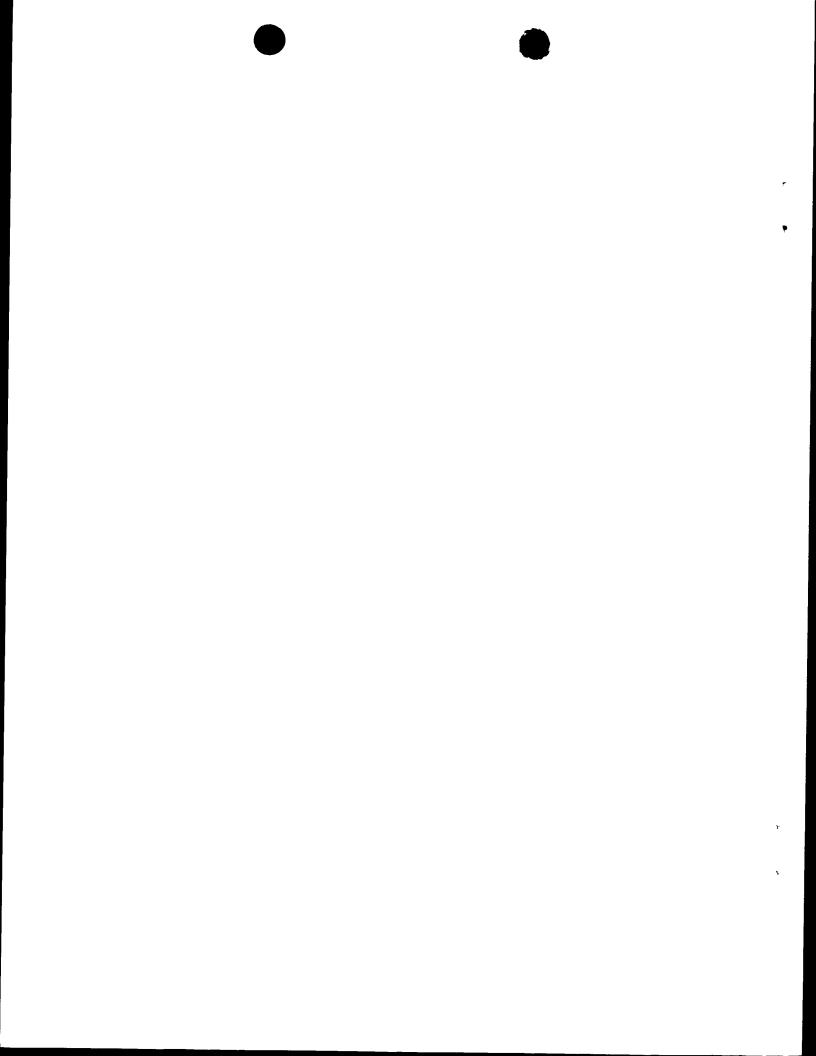
```
<110> BASF AG, Forschungszentrum Jülich GmbH
<120> Ein- oder mehrzellige Organismen zur Herstellung von
      Riboflavin
<130> FZJ9909PCT
<140> PCT/EP00/07370
<141> 2000-07-31
<160> 2
 <170> PatentIn Ver. 2.1
 <210> 1
 <211> 2321
 <212> DNA
 <213> Pilz der Spezies Ashbya Gossypii
 <220>
 <221> CDS
 <222> (718)..(1266)
 ctgcagcaaa tcgaggtgat cgccaacgag gtggacgtgc ggcaggacgg gacctggtgc 60
 atccggtacc gcgacgagtc cgagcacggg cacgacaagt cgcggtcgat cgcggcgtgc 120
 aagcagcgct ggcaacacct cgagcccgcg ccggtgtatt tctactgcgg cgatgggatc 180
  agegacetga gegetgegaa ggaatgegae etgetgtttg egaagagtgg caaggacetg 240
  atctccttct gcaagaagca ggacgttccg ttccgcgagt tcaacacttt tgacgatgtg 300
  ctgagegegg teaagegegt ggtggeggge gaggeetetg teaeggaaet eeagggggge 360
  teegetgegt aageaetgte tgeateagtg acettggegg tagetgegat ttgtaactae 420
  ctacgtaatt agtcctgctc gcgctgcggt ccagtgctag gcacgcccca catgaaaggc 480
  agccgtaagc aattagtaac ggcctagtac ggctccgatg tatgtgctag cacatgacag 540
  cccaacgggt tgagaagtcc ggctcgaatc atttccgcgc cgagtgggtc gtgggtggag 600
   ccgcccgacc ccttgtcagc gcgggcagtt ggatataagg cagtggttgt agcaaaagtg 660
   agXtgcgtgca tttcacgaag ccgagcgcaa caacgcacag acatcagtaa gcagct atg 720
                                                                    1
   ggc aag gtg aag gta caa caa ccc atc gtc gag atg gac ggc gac gaa
                                                                      768
   Gly Lys Val Lys Val Gln Gln Pro Ile Val Glu Met Asp Gly Asp Glu
   cag acg cgg atc atc tgg cac ttg atc aag gat cag ctc atc ttc ccc
                                                                      816
   Gln Thr Arg Ile Ile Trp His Leu Ile Lys Asp Gln Leu Ile Phe Pro
```

25

20



2
tac ttg gac gtg gac ttg aag tac tac gat ctt tcc att gag aac agg 864  Tyr Leu Asp Val Asp Leu Lys Tyr Tyr Asp Leu Ser Ile Glu Asn Arg  45
gat gcc acc gag gac cgc gtg act gtg gag tct gcg gag gcg acc ctc 912 Asp Ala Thr Glu Asp Arg Val Thr Val Glu Ser Ala Glu Ala Thr Leu 60 65
aag tac ggc gtt gcc gtc aag tgt gcg att att acc ccg gac gag gcg 960 Lys Tyr Gly Val Ala Val Lys Cys Ala Ile Ile Thr Pro Asp Glu Ala 75 80
cgt gtc gag gag ttc ggg ctc aag gag atg tgg aag tct ccc aac ggg 1008 Arg Val Glu Phe Gly Leu Lys Glu Met Trp Lys Ser Pro Asn Gly 85 90 95
acc atc cgg aac atc ctc ggc ggg acc gtc ttc aga gag ccc att att 1056  Thr Ile Arg Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile Ile  100 105 110
atc cca agg atc ccc aga ctg gtg ccc ggc tgg aac gag ccg atc att 1104  Ile Pro Arg Ile Pro Arg Leu Val Pro Gly Trp Asn Glu Pro Ile Ile  115 120 125
gtc ggc aga cac gcg ttt ggg gac cag tac aag gcg acc gac gtt gtc 1152 Val Gly Arg His Ala Phe Gly Asp Gln Tyr Lys Ala Thr Asp Val Val 130 145
att cca ggc gag ggc acg ttg aag ctg gtc ttt gaa agc aag gac ggg 1200  Ile Pro Gly Glu Gly Thr Leu Lys Leu Val Phe Glu Ser Lys Asp Gly  150  150
gac aag too aag aat ott gac otg gag tto ttt gaa tac ooc aag gat 1248 Asp Lys Ser Lys Asn Leu Asp Leu Glu Phe Phe Glu Tyr Pro Lys Asp 165 170 175
ggc ggt gtt gcc atg acc atgtactaca ccaccgactc gatcaccggc 1296 Gly Gly Val Ala Met Thr 180
tttgccaagt cgagcttcga gttggcgttg caaagaaaga tgccgctata ttcgacaacg 1356
aagaacacga tottgaagaa gtacgacggo aagtttaagg atattttcga gggcatgtac 1416
ccagcggagt acaaggagaa gtttgaggct gctggcatct ggtatgaaca cagactgatt 1476
gacqatatgg ttgcgcagat gttgaagtcc aagggcggct tcatcattgc catgaagaac 1536
tacgatggtg atgtgcagtc ggacatcgtc gcccagggct tcgggtcttt gggtctcatg 1596
acttetgtte ttgtgtetee agatggaaag acettegaga gtgaggeege acatggeaet 1656
gtcaccegge actacagaca gcaccagcag ggcaaggaaa catecaccaa ctetattgee 1716
totatttttg cotggatgog oggtattata cacagaggta aggtogaogg taccocagat 1776
gtegtgaagt teggegagtt gttggagaag teeaceetgg acaeggtgea ggaggacate 1836
atgaccaagg acctagcgtt gattttgggc aagaccgaca gagccagcta tgttaccacg 1896



gaagagttta tcacagcagt agcgaaccgc ttagcgatgg ctacaagcgt cttttttgtg 1956 aataagaaaa agcaagcaaa attatagcct aggctgcctg tagcgtctat ttattactag 2016 tetageatat etageacaag aatatagata etgageeate egeceaggat tacagteagg 2076 attocaactt gtaaacctcc ggtggtgcgc actcgccgca aattaggtga gcttgccatt 2136 agtcatccga ggcgcagaat gagtagggtt tatagtaaac ccgggtgctg taacaccaga 2196 teccaetttt cetggeacag tatttttgee gacaaeggea etgetaaeeg ttteteaaet 2256 acgcgcaata atgtaggtcg cacggtccga tgaaaactaa tgcgcagtag catgacatgg 2316 2321 aattc

<210> 2

<211> 183

<212> PRT

<213> Pilz der Spezies Ashbya Gossypii

<400> 2

Met Gly Lys Val Lys Val Gln Gln Pro Ile Val Glu Met Asp Gly Asp

Glu Gln Thr Arg Ile Ile Trp His Leu Ile Lys Asp Gln Leu Ile Phe

Pro Tyr Leu Asp Val Asp Leu Lys Tyr Tyr Asp Leu Ser Ile Glu Asn

Arg Asp Ala Thr Glu Asp Arg Val Thr Val Glu Ser Ala Glu Ala Thr 50

Leu Lys Tyr Gly Val Ala Val Lys Cys Ala Ile Ile Thr Pro Asp Glu

Ala Arg Val Glu Glu Phe Gly Leu Lys Glu Met Trp Lys Ser Pro Asn

Gly Thr Ile Arg Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile

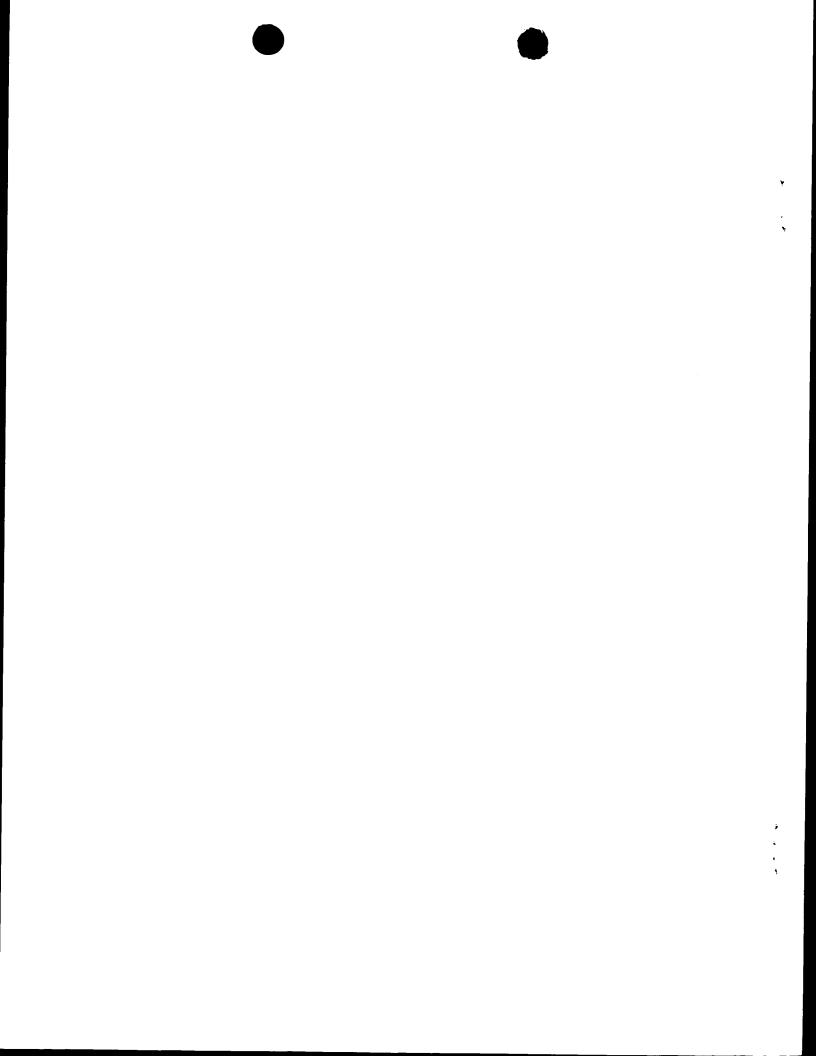
Ile Ile Pro Arg Ile Pro Arg Leu Val Pro Gly Trp Asn Glu Pro Ile

Ile Val Gly Arg His Ala Phe Gly Asp Gln Tyr Lys Ala Thr Asp Val

Val Ile Pro Gly Glu Gly Thr Leu Lys Leu Val Phe Glu Ser Lys Asp 150

Gly Asp Lys Ser Lys Asn Leu Asp Leu Glu Phe Phe Glu Tyr Pro Lys 170 165

Asp Gly Gly Val Ala Met Thr





(19) Weltorganisation für geistiges Eigentum Internationales Büro



# 

(43) Internationales Veröffentlichungsdatum 15. Februar 2001 (15.02.2001)

PCT

### (10) Internationale Veröffentlichungsnummer WO 01/11052 A3

- C12N 15/53, (51) Internationale Patentklassifikation7: 15/80, 9/04, 1/15, C12P 25/00 // (C12N 1/15, C12R 1:645)
- (21) Internationales Aktenzeichen:

PCT/EP00/07370

(22) Internationales Anmeldedatum:

31. Juli 2000 (31.07.2000)

(25) Einreichungssprache:

Deutsch

(26) Veröffentlichungssprache:

Deutsch

(30) Angaben zur Priorität:

199 37 548.8

9. August 1999 (09.08.1999) DE

- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; D-67056 Ludwigshafen (DE). FORSCHUNGSZEN-TRUM JÜLICH GMBH [DE/DE]; D-52425 Jülich (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): ALTHÖFER, Henning [DE/DE]; Mainstrasse 12, D-67117 Limburgerhof (DE). ZELDER, Oskar [DE/DE]; Rossmarktstrasse 27, D-67346 Speyer (DE). REVUELTA DOVAL, Jose L. [ES/ES]; Grillo 11 4E, E-37001 Salamanca (ES). SANTOS GARCIA, Maria Angeles [ES/ES]; C/Escuelas, 1-5, E-37001 Salamanca (ES). SAHM, Hermann [DE/DE]; Wendelinusstrasse 71, D-52428 Jülich (DE). STAHMANN, Klaus-Peter [DE/DE]; Wilhelmstrasse

- 20, D-52428 Jülich (DE). MAETING, Ines [DE/DE]; Wiesenstr. 3, D-52428 Jülich (DE).
- (74) Anwalt: FITZNER, Uwe; Lintorfer Strasse 10, D-40878 Ratingen (DE).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Veröffentlicht:

- Mit internationalem Recherchenbericht.
- (88) Veröffentlichungsdatum des internationalen 5. Juli 2001 Recherchenberichts:

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: MONOCELLULAR OR MULTICELLULAR ORGANISMS FOR THE PRODUCTION OF RIBOFLAVIN

(54) Bezeichnung: EIN- ODER MEHRZELLIGE ORGANISMEN ZUR HERSTELLUNG VON RIBOFLAVIN

(57) Abstract: The invention relates to a monocellular or multicellular organism, especially a microorganism, for biotechnological production of riboflavin, whereby the enzymatic activity thereof with respect to NAD(P)H formation is higher than that of a wild type of species Ashbya gossypii ATCC10895.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft einen ein- oder mehrzelligen Organismus, insbesondere Mikroorganismus, zur biotechnischen Herstellung von Riboflavin, wobei dessen Enzymaktivität bezüglich der NAD(P)H Bildung höher ist als derjenige eines Wildtyps der Species Ashbya gossypii ATCC10895.



7 **\*** 

# INTERNATIONAL SEARCH REPORT



A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/53 C12N15/80 //(C12N1/15,C12R1:645)

C12N9/04

C12N1/15

C12P25/00

According to International Patent Classification (IPC) or to both national classification and IPC

#### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

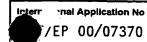
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EPO-Internal, MEDLINE

Category °	NTS CONSIDERED TO BE RELEVANT  Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X .	HENKE BIRGIT ET AL: "IDP3 encodes a peroxisomal NADP-dependent isocitrate dehydrogenase required for the beta-oxidation of unsaturated fatty acids."  JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 6, 6 February 1998 (1998-02-06), pages 3702-3711, XP002157778  ISSN: 0021-9258 page 3705; figure 3 page 3709, column 2 -page 3710  -/	7-10

Further documents are tisted in the continuation of box C.	X Patent latting thembers do noted in
Special categories of cited documents:      A' document defining the general state of the art which is not considered to be of particular relevance.	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
<ul> <li>"E" earlier document but published on or after the international filing date</li> <li>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</li> <li>"O" document referring to an oral disclosure, use, exhibition or other means</li> <li>"P" document published prior to the international filing date but later than the priority date claimed</li> </ul>	<ul> <li>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</li> <li>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</li> <li>"&amp;" document member of the same patent family</li> </ul>
Date of the actual completion of the international search	Date of mailing of the international search report
19 January 2001	01/02/2001
Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Blanco Urgoiti, B

# INTERNATIONAL SEARCH REPORT



		/EF U	0/07370		
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT					
Category °	Citation of document, with indication, where appropriate, of the relevant passages		Relevant to claim No.		
Α	WO 97 03208 A (BASF AG; KERNFORSCHUNGSANLAGE JUELICH (DE); KAESLER BRUNO (DE); SA) 30 January 1997 (1997-01-30) cited in the application the whole document				
A	EP 0 927 761 A (BASF AG) 7 July 1999 (1999-07-07) the whole document				
A	EP 0 405 370 A (HOFFMANN LA ROCHE) 2 January 1991 (1991-01-02) the whole document				

# INTERNATIONAL SEARCH REPORT

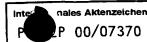
on patent family members

Internal Application No FP 00/07370

Patent document cited in search repor	t	Publication date		tent family ember(s)	Publication date
WO 9703208	Α	30-01-1997	DE DE CA CN EP JP US	19525281 C 19545468 A 2223877 A 1193356 A 0839211 A 11509409 T 5976844 A	04-04-1996 21-08-1997 30-01-1997 16-09-1998 06-05-1998 24-08-1999 02-11-1999
EP 0927761	Α	07-07-1999	CN WO EP JP	1227870 A 9933993 A 1040193 A 11243975 A	08-09-1999 08-07-1999 04-10-2000 14-09-1999
EP 0405370	A	02-01-1991	AT CN DE EP JP JP US US	195971 T 1049185 A 69033616 D 1001026 A 3117489 A 10066562 A 5925538 A 5837528 A	15-09-2000 13-02-1991 05-10-2000 17-05-2000 20-05-1991 10-03-1998 20-07-1999 17-11-1998

ĺ

# INTERNATIONALER RECHERCHENBERICHT



KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES PK 7 C12N15/53 C12N15/80 C12P25/00 C12N1/15 C12N9/04

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

#### B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole ) C12P IPK 7 C12N

//(C12N1/15,C12R1:645)

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtt. verwendete Suchbegriffe)

BIOSIS, EPO-Internal, MEDLINE

ategorie°	Bezeichnung der Veröffentlichung, soweit erfordertich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	HENKE BIRGIT ET AL: "IDP3 encodes a peroxisomal NADP-dependent isocitrate dehydrogenase required for the beta-oxidation of unsaturated fatty acids."  JOURNAL OF BIOLOGICAL CHEMISTRY, Bd. 273, Nr. 6, 6. Februar 1998 (1998-02-06), Seiten 3702-3711, XP002157778 ISSN: 0021-9258 Seite 3705; Abbildung 3 Seite 3709, Spalte 2 -Seite 3710  -/	7-10

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu	X Siehe Anhang Patentfamilie
enthethien	*T* Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum
Besondere Kategorien von angegebenen Veröffentlichungen :	
*A* Veröffentlichung, die den altgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist	oder dem Frontatsdatung Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden
*E* älteres Dokument, das jedoch erst am oder nach dem internationalen	Theorie angegeben ist

Anmeldedatum veröffentlicht worden ist

\*L\* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft er-scheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie

\*O\* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht \*P\* Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach

dem beanspruchten Prioritätsdatum veröffentlicht worden ist

Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

\*&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Absendedatum des internationalen Recherchenberichts

Datum des Abschlusses der internationalen Recherche

19. Januar 2001

Name und Postanschrift der Internationalen Recherchenbehörde

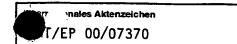
Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016

01/02/2001

Bevollmächtigter Bediensteter

Blanco Urgoiti, B

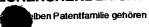
# INTERNATIONALE RECHERCHENBERICHT



		00/07370						
	(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN							
Kategorie®	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.						
Α	WO 97 03208 A (BASF AG; KERNFORSCHUNGSANLAGE JUELICH (DE); KAESLER BRUNO (DE); SA) 30. Januar 1997 (1997-01-30) in der Anmeldung erwähnt das ganze Dokument							
A	EP 0 927 761 A (BASF AG) 7. Juli 1999 (1999-07-07) das ganze Dokument							
A	EP 0 405 370 A (HOFFMANN LA ROCHE) 2. Januar 1991 (1991-01-02) das ganze Dokument							

## INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichung..., die



pales Aktenzeichen P 00/07370

	Im Recherchenbericht angeführtes Patentdokument WO 9703208 A		Mitglied( Patentf		Datum der Veröffentlichung
WO 9703208			DE 199 CA 22 CN 11 EP 08 JP 119	525281 C 545468 A 223877 A 193356 A 839211 A 509409 T 976844 A	04-04-1996 21-08-1997 30-01-1997 16-09-1998 06-05-1998 24-08-1999 02-11-1999
EP 0927761	Α	07-07-1999	WO 9'	227870 A 933993 A 040193 A 243975 A	08-09-1999 08-07-1999 04-10-2000 14-09-1999
EP 0405370	A	02-01-1991	CN 1 DE 69 EP 1 JP 3 JP 10 US 5	195971 T 049185 A 033616 D 001026 A 117489 A 066562 A 925538 A	15-09-2000 13-02-1991 05-10-2000 17-05-2000 20-05-1991 10-03-1998 20-07-1999 17-11-1998

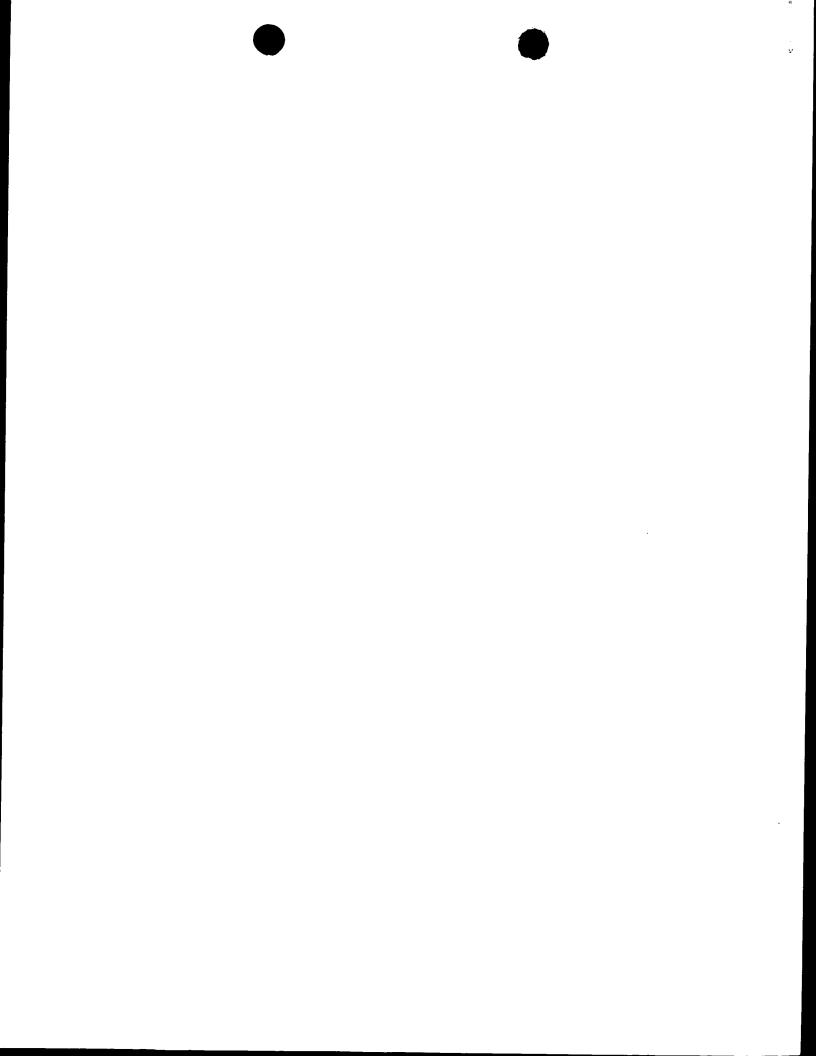
# VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWE

PCT

REC'D 3 0 NOV 2001 INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT FOT

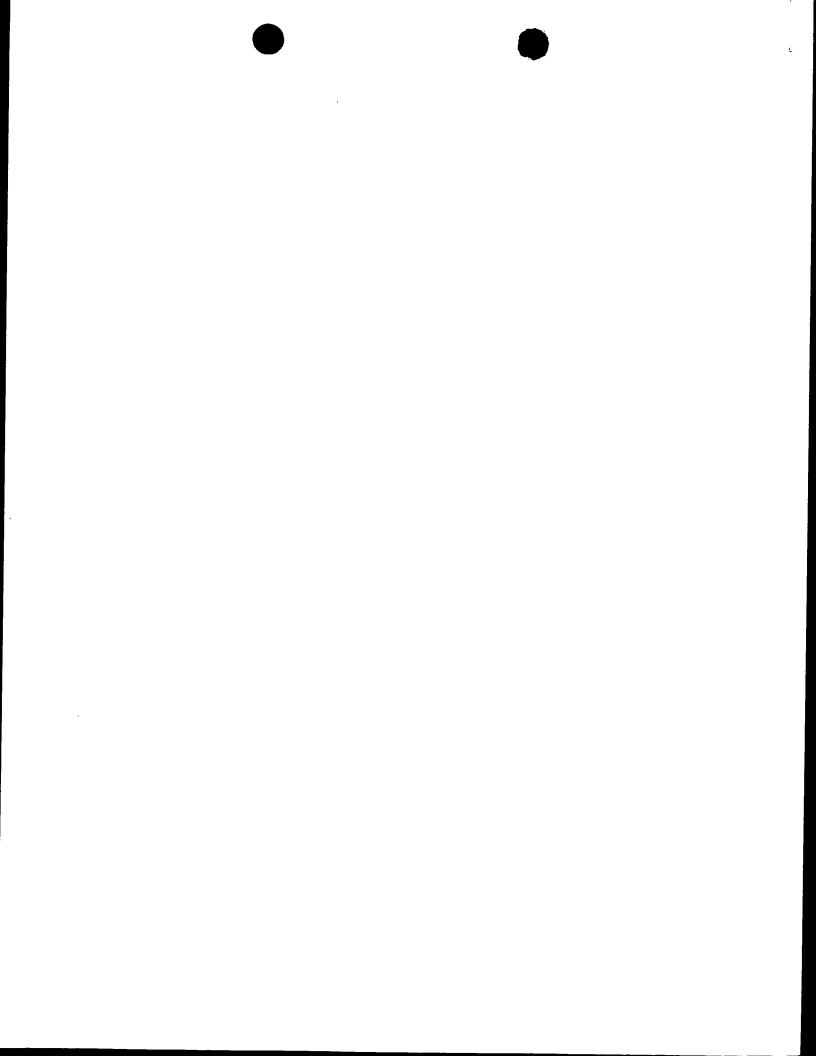
(Artikel 36 und Regel 70 PCT)

	(Artikel 66 and 1 legel :	
Aktenzeichen des Anmelders oder Anwalts	WEITERES VORGEHEN vo	ehe Mitteilung über die Übersendung des internationalen orläufigen Prüfungsberichts (Formblatt PCT/IPEA/416)
FZJ 9909 PCT		
Internationales Aktenzeichen	Internationales Anmeldedatum(Tag/Mo	i I
PCT/EP00/07370	31/07/2000	09/08/1999
Internationale Patentklassifikation (IPK) oder	nationale Klassifikation und IPK	
C12N15/53		
Anmelder		
BASF AKTIENGESELLSCHAFT et a	al.	
		Different beguttregten
Dieser internationale vorläufige Prü     Behörde erstellt und wird dem Anm	fungsbericht wurde von der mit der elder gemäß Artikel 36 übermittelt.	internationalen vorläufigen Prüfung beauftragten
2. Dieser BERICHT umfaßt insgesam	t 6 Blätter einschließlich dieses Dec	ckblatts.
M	ANII AGEN hoi: dahoi handelt es sic	h um Blätter mit Beschreibungen, Ansprüchen
Lundlador Zajahnungan, dia gai	indert wurden und diesem Bericht z	uarunde liegen, und/oder Blatter mit voi diesei
Behörde vorgenommenen Ber	ichtigungen (siehe Regel 70.16 und	Abschnitt 607 der Verwaltungsrichtlinien zum PCT).
Diese Anlagen umfassen insgesan	nt 3 Blätter	
Diese Arilagen umlassen magesam	N O Diano.	
3. Dieser Bericht enthält Angaben zu	folgenden Punkten:	
_		
I ⊠ Grundlage des Bericht	S	
II □ Priorität	Cutachtone über Neuheit erfinderi	sche Tätigkeit und gewerbliche Anwendbarkeit
		John Fallghon and government
IV Mangelnde Einheitlich	keit der Emindung	r Neuheit, der erfinderischen Tätigkeit und der
gewerblichen Anwend	barkeit; Unterlagen und Erklärungei	n zur Stützung dieser Feststellung
VI   Bestimmte angeführte		
	r internationalen Anmeldung	
VIII   Bestimmte Bemerkung	gen zur internationalen Anmeldung	
	·	
Datum der Einreichung des Antrags	Datum der	Fertigstellung dieses Berichts
01/03/2001	28.11.200	1
Name und Postanschrift der mit der internat	ionalen vorläufigen Bevollmäc	htigter Bediensteter
Prüfung beauftragten Behörde:	[	\$ 1 m
Europäisches Patentamt D-80298 München	A. M. Me	erlos ( 🖁 🐠 🐉
Tel. +49 89 2399 - 0 Tx: 5236	56 epmu d	Brann Street
Fax: +49 89 2399 - 4465	Tel. Nr. +4	9 89 2399 8559



#### I. Grundlage des Berichts

		lulage des Derio.				
1.	Auffc einge	andamma naab Art	ndteile der internationalen Ar ikel 14 hin vorgelegt wurden, ihm nicht beigefügt, weil sie ko n:	aelten im Kahm	en dieses Berichts als	s ursprungiion
	1-8		ursprüngliche Fassung			
	Pate	ntansprüche, Nr	:			
	1-20		eingegangen am	13/09/2001	mit Schreiben vom	10/09/2001
	Zeic	hnungen, Blätter	r:			
	1-11		ursprüngliche Fassung			
2.	die i	nternationale Ann	che: Alle vorstehend genannte neldung eingereicht worden is ichts anderes angegeben ist.	en Bestandteile it, zur Verfügung	standen der Behörde 3 oder wurden in diese	in der Sprache, in der er eingereicht, sofern
	Die eing	Bestandteile stand Jereicht; dabei har	den der Behörde in der Sprac ndelt es sich um	he: zur Verfüg	ung bzw. wurden in di	eser Sprache
		die Sprache der U Regel 23.1(b)).	Übersetzung, die für die Zwec	cke der internati	onalen Recherche ein	gereicht worden ist (nach
		die Veröffentlicht	ungssprache der international	en Anmeldung (	nach Regel 48.3(b)).	
		die Sprache der ist (nach Regel 5	Übersetzung, die für die Zwed 5.2 und/oder 55.3).	cke der internati	onalen vorläufigen Pr	üfung eingereicht worden
3	. Hins inte	sichtlich der in der rnationale vorläufi	r internationalen Anmeldung o ige Prüfung auf der Grundlage	offenbarten <b>Nuc</b> e des Sequenzp	l <b>eotid- und/oder Ami</b> protokolls durchgeführ	nosäuresequenz ist die t worden, das:
		in der internation	alen Anmeldung in schriftliche	er Form enthalte	en ist.	
			er internationalen Anmeldung			nt worden ist.
			nachträglich in schriftlicher Fo			
			nachträglich in computerlesb			
		Die Erklärung, da Offenbarungsge	aß das nachträglich eingereic halt der internationalen Anme	hte schriftliche S Idung im Anmel	Sequenzprotokoll nich dezeitpunkt hinausgel	nt, wurde vorgelegt.
		Die Erklärung, d	aß die in computerlesbarer Fo oll entsprechen, wurde vorgele	orm erfassten In		
_	. Auf	farund der Änderu	ingen sind folgende Unterlage	en fortgefallen:		

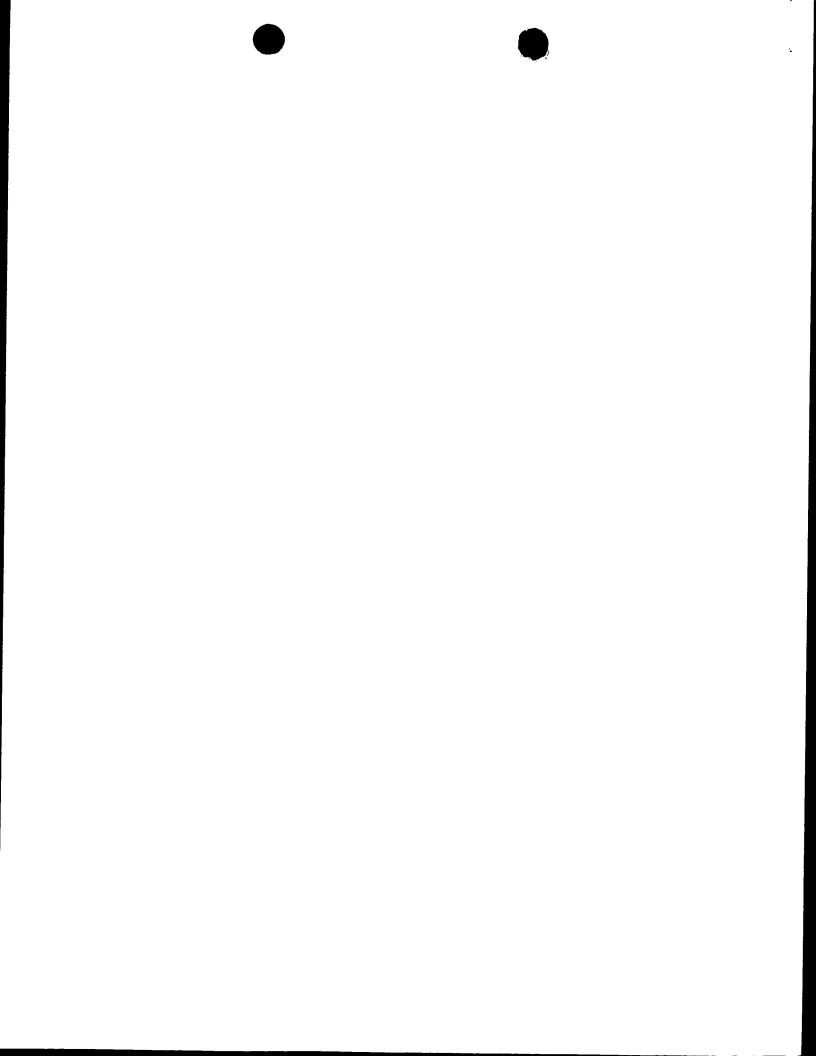


		Beschreibung, Ansprüche, Zeichnungen,	Seiten: Nr.: Blatt:			
5.		Dieser Bericht ist ohr angegebenen Gründ eingereichten Fassu	en nach Auffa	assur	ng der Behörd	n) der Änderungen erstellt worden, da diese aus den e über den Offenbarungsgehalt in der ursprünglich
		(Auf Ersatzblätter, di beizufügen).	e solche Änd	lerung	gen enthalten,	ist unter Punkt 1 hinzuweisen;sie sind diesem Bericht
6.	Etw	aige zusätzliche Bem	erkungen:			
۷.	Beç gev	gründete Feststellun verblichen Anwendb	g nach Artik earkeit; Untei	el 356 rlage	(2) hinsichtlic n und Erklärt	ch der Neuheit, der erfinderischen Tätigkeit und de Ingen zur Stützung dieser Feststellung
1.	Fes	ststellung				
	Neu	uheit (N)			Ansprüche Ansprüche	3-5, 7, 8; 9-14 teilweise; 15-17; 18-20 teilweise 1, 2, 6; 9-14, teilweise; 18-20 teilweise
	Erfi	nderische Tätigkeit (E	<del>-</del> · /	Ja: Vein:	Ansprüche Ansprüche	3-5 teilweise, 7, 8; 9-14, teilweise; 18-20 teilweise 1, 2; 3-5 teilweise; 6; 9-14 teilweise; 15-17; 18-20 teilweise
	Ge	werbliche Anwendbar		Ja: Nein:	Ansprüche Ansprüche	1-20

2. Unterlagen und Erklärungen siehe Beiblatt

# VIII. Bestimmte Bemerkungen zur internationalen Anmeldung

Zur Klarheit der Patentansprüche, der Beschreibung und der Zeichnungen oder zu der Frage, ob die Ansprüche in vollem Umfang durch die Beschreibung gestützt werden, ist folgendes zu bemerken: siehe Beiblatt

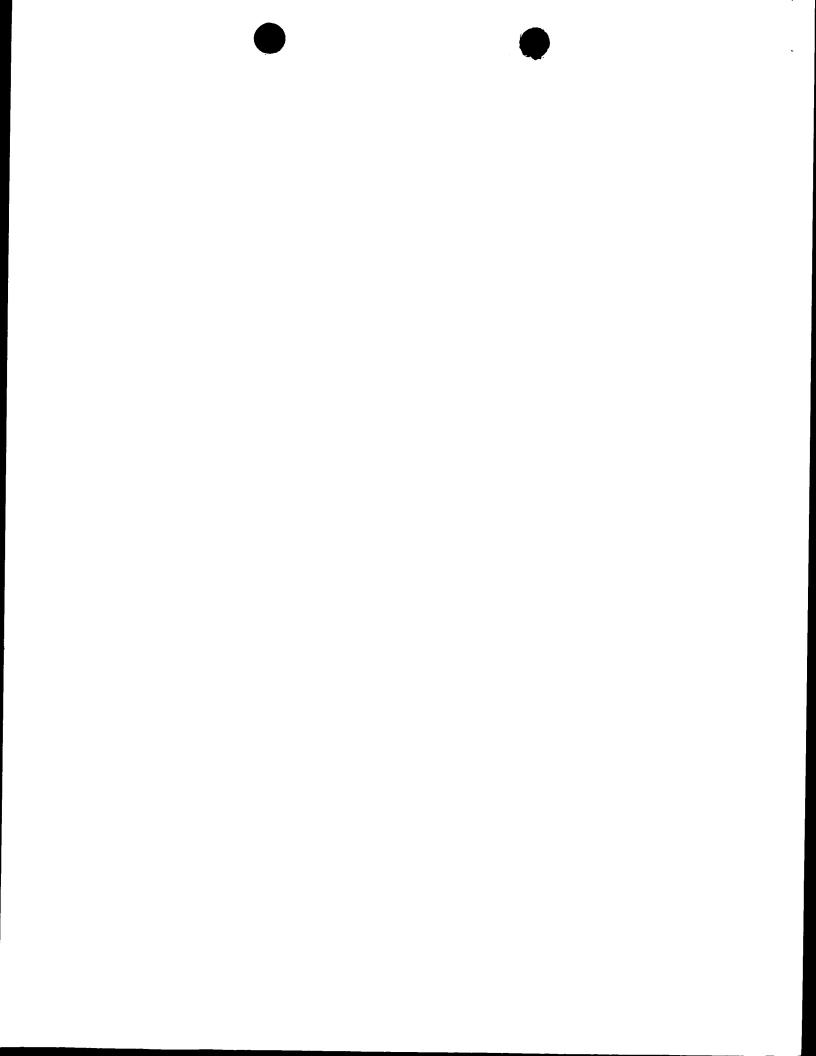


### 1. Art. 34 (2,b) PCT

Die geänderten Ansprüche 1-20 scheinen den Anforderungen des Art. 34(2,b) PCT zu entsprechen. Es sei in diesem Zusammenhang aber darauf hingewiesen, dass offensichtlich die ursprünglichen Anmeldungsunterlagen, auf die zur Stützung der eingeführten Änderungen Bezug genommen wurde, nicht mit den der Behörde vorliegenden Anmeldungsunterlagen übereinstimmen. Die Anmeldung wie ursprünglich eingereicht, enthält eine Beschreibung von acht Seiten. Mit Hinblick auf die geänderten Ansprüche 6, 7 und 8 wurde aber beispielsweise auf eine Seite 11, Zeilen 29-34 Bezug genommen. Ebenso befindet sich auf Seite 7, Zeilen 19-30 keine sinngemässe Offenbarung für den geänderten Anspruch 9.

## 2. Art. 33(2), (3) PCT

Ein Teil der geänderten Ansprüche 1-20 scheint nach wie vor nicht den Anforderungen des Art. 33(2), (3) PCT zu entsprechen. Dokument "Henke et al., J. of Biol Chem., Bd. 273, Nr. 6, 1998, 3702-3711", beschreibt die Identifizierung einer peroxisomalen Isocitrat-Dehydrogenase (Idp3) und des sie kodierenden Gens aus S. cerevisiae, sowie dessen rekombinante Expression in entsprechend transformierten E. coli Zellen (s. Fig. 3, S. 3706, linke Spalte, zweiter Absatz und rechte Spalte, einschl. zweiter Absatz, Seite 3709, rechte Spalte erster Absatz). Die transformierten E. coli Zellen zeigten einen massiven Anstieg der Isocitratedehydrogenase-Aktivität. Anspruch 1 bezieht sich auf einen Mikroorganismus zur biotechnischen Herstellung von Riboflavin, der dadurch gekennzeichnet ist, dass er eine Aktivität eines NAD(P)H bildenden Enzyms aufweist, die höher ist als diejenige eines Wildtyps der Species Ashbya gossypii ATCC 10895 (s. auch unter Punkt 3.). Der beabsichtigte Verwendungszweck, also die biotechnische Herstellung von Riboflavin, ist als optionales und daher nicht limitierendes Merkmal anzusehen, solange nicht den Mikroorganismus entsprechend kennzeichnende spezifische (d.h. neuheitsgebende) und speziell für diesen Zweck benötigte Merkmale aufgeführt sind. Daher wäre E. coli, welches rib Gene enthält (s. beispielsweise E0 0 405 370, Seite 3, Zeilen 25-30) gleichermassen für die Biosynthese von Riboflavin geeignet, auch wenn dies von Henke et al. nicht nachgewiesen wurde, da hier lediglich die Verifizierung der klonierten Idp3 im Vordergrund stand. Zudem sei darauf hingewiesen, dass sich Anspruch 1 auf einen Mikroorganismus bezieht, der eine erhöhte Aktivität



3

bezüglich eines NAD(P)H bildenden Enzyms aufweist. Zur Höhe des tatsächlich in der Zelle gebildeten NAD(P)H-Gehalts wird auch hier nicht Bezug genommen. Anspruch 1 wird daher nicht als neu und erfinderisch im Sinne von Art. 33(2), (3) PCT angesehen. Dies gilt ebenso für den abhängigen Anspruch 2 und die Ansprüche 11, 12, 19 und 20, soweit sie sich direkt oder indirekt auf Anspruch 6 rückbeziehen (s. unten).

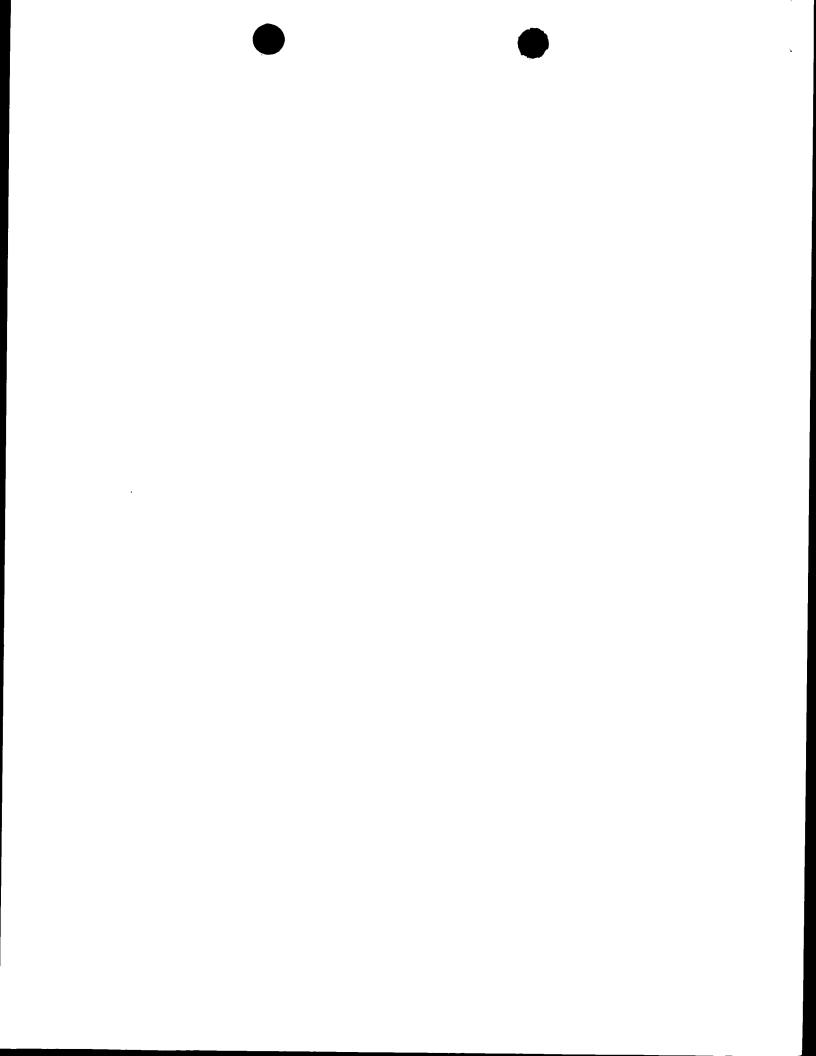
Anspruch 6 wird nicht als neu und erfinderisch angesehen, da das aus Henke et al. bekannte ldp3-Gen zu 65% identisch ist mit der in der gegenwärtigen Anmeldung beschriebenen DNA-Sequenz des ldp3-Gens und somit unter den Begriff "Allelvariation" fällt (s. Definition des Begriffes Allelvariation auf Seite 6, Zeilen 3-6). Dies gilt ebenfalls für die Ansprüche 9 und 10.

Die Ansprüche 3-5 scheinen sich auf einen neuen und soweit sie sich auf Anspruch 2 rückbeziehen auch auf einen erfinderischen Gegenstand zu beziehen. Dies gilt auch für die direkt abhängigen Ansprüche 14, 18. Ansprüche 7, 8 sowie davon direkt oder indirekt abhängige Ansprüche 9-14 und 18 können ebenfalls als neu und erfinderisch angesehen werden.

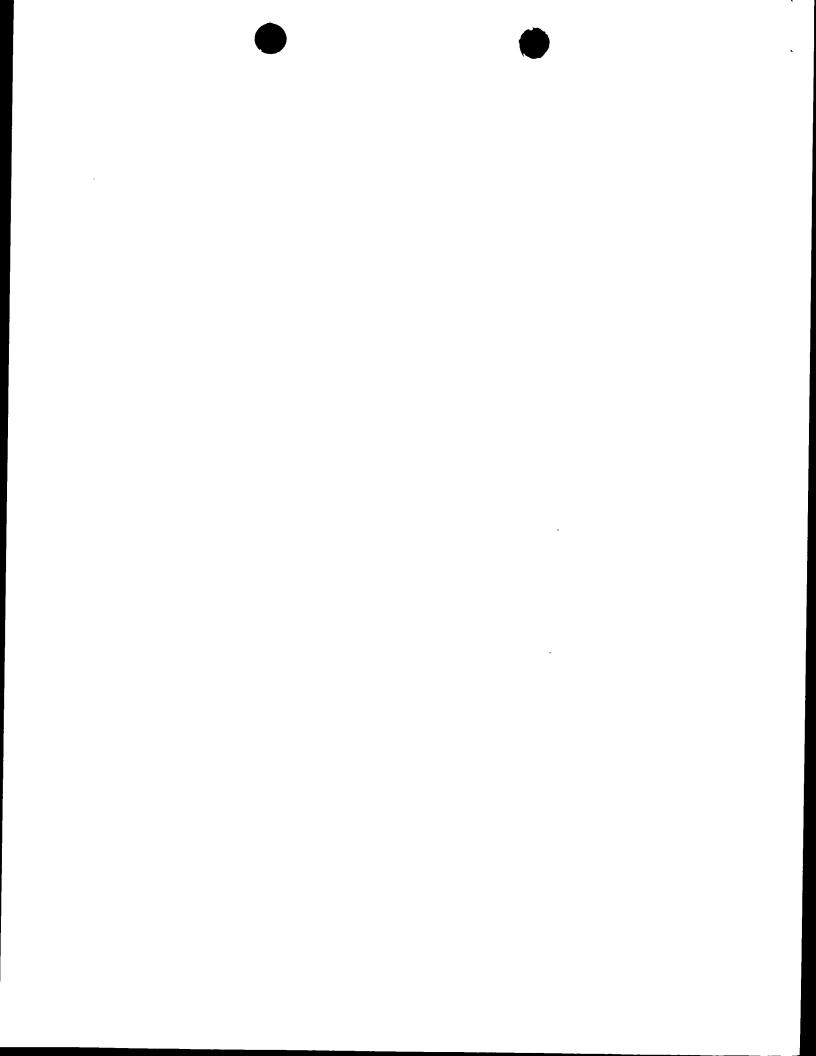
#### Art. 5/6 PCT 3.

Der gegenwärtigen Erfindung liegt die Entdeckung zugrunde, dass mit einer verstärkten Isocitrat-dehydrogenase-Expression im Pilz Ashbya ATCC 10895 auch ein deutlicher Anstieg der Riboflavinbildung einhergeht. Allerdings lässt sich aus dieser Beobachtung kein allgemein gültiges Prinzip ableiten, wonach der Anstieg irgendeines anderen NAD(P)H bildenden Enzyms zum gleichen Ergebnis führt. Mit Hinblick auf die auf Idp3 limitierte Beschreibung, sind die Ansprüche, die lediglich vage auf ein NAD(P)H bildendes Enzym gerichtet sind, zu breit und daher nicht mit den Erfordernissen des Art. 5 vereinbar.

Desweiteren sei darauf hingewiesen, dass ohne spezifischen Vergleichswert oder klare Definition des Vergleichswerts (bspw. Idp3-Gehalt unter spezifischen Wachstumbedingungen) oder Angabe eines präzisen, strukturellen Merkmals (bspw. verstärkte Expression oder erhöhte Kopienzahl des Idp3-Gens, die beide eine erhöhte Aktivität zur Folge haben, s. auch Anspruch 11), Anspruch 1 (auch Ansprüche 2-3) unklar ist (Art. 6 PCT), da aus der derzeitigen Formulierung nicht eindeutig hervorgeht, aufgrund welcher Umstände die Aktivität erhöht ist. So könnten auch lediglich geänderte Kulturbedingungen eine erhöhte Idp3-Aktivität bewirken. Zudem scheint ein Vergleich nur zwischen zwei Mikroorganismen der

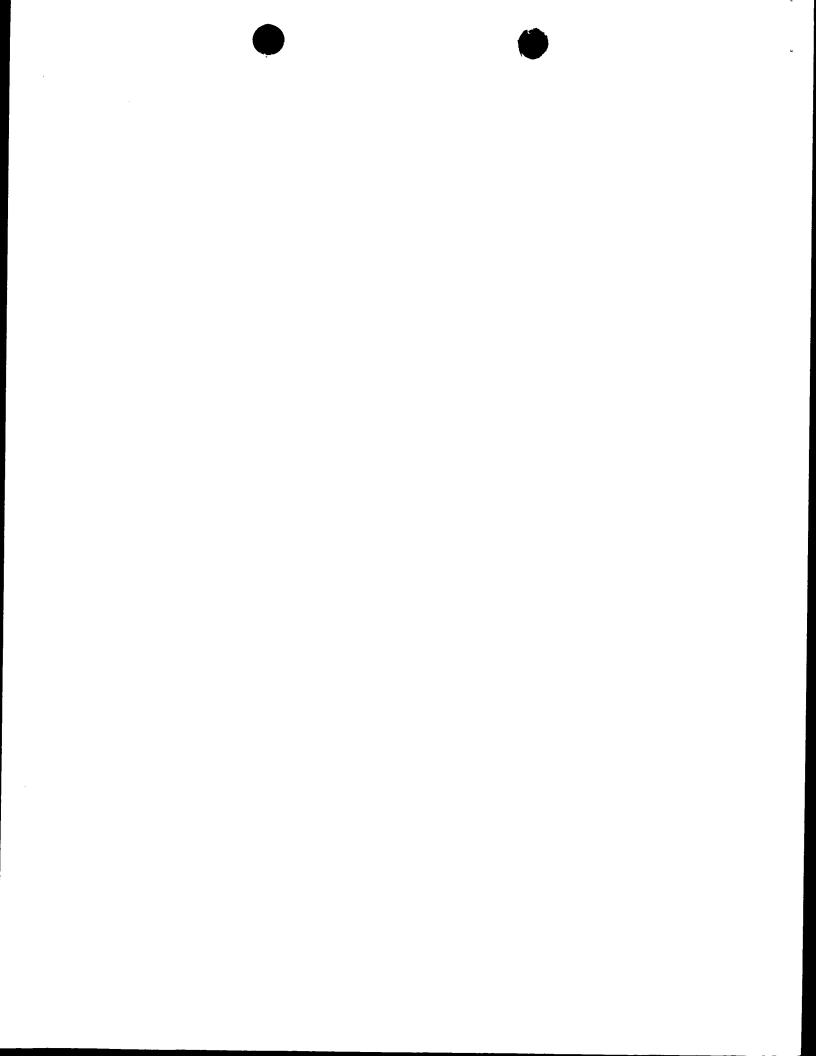


4. Schliesslich sei der Anmelder darauf hingewiesen, dass sich das PCT und Euro-Verfahren unter anderem darin unterscheiden, dass in der PCT-Phase als letzte Aktion lediglich ein vorläufiger Prüfungsbericht erstellt wird, aber keine bindende Entscheidung getroffen wird, sodass ein Recht auf rechtliches Gehör nicht verletzt werden kann.



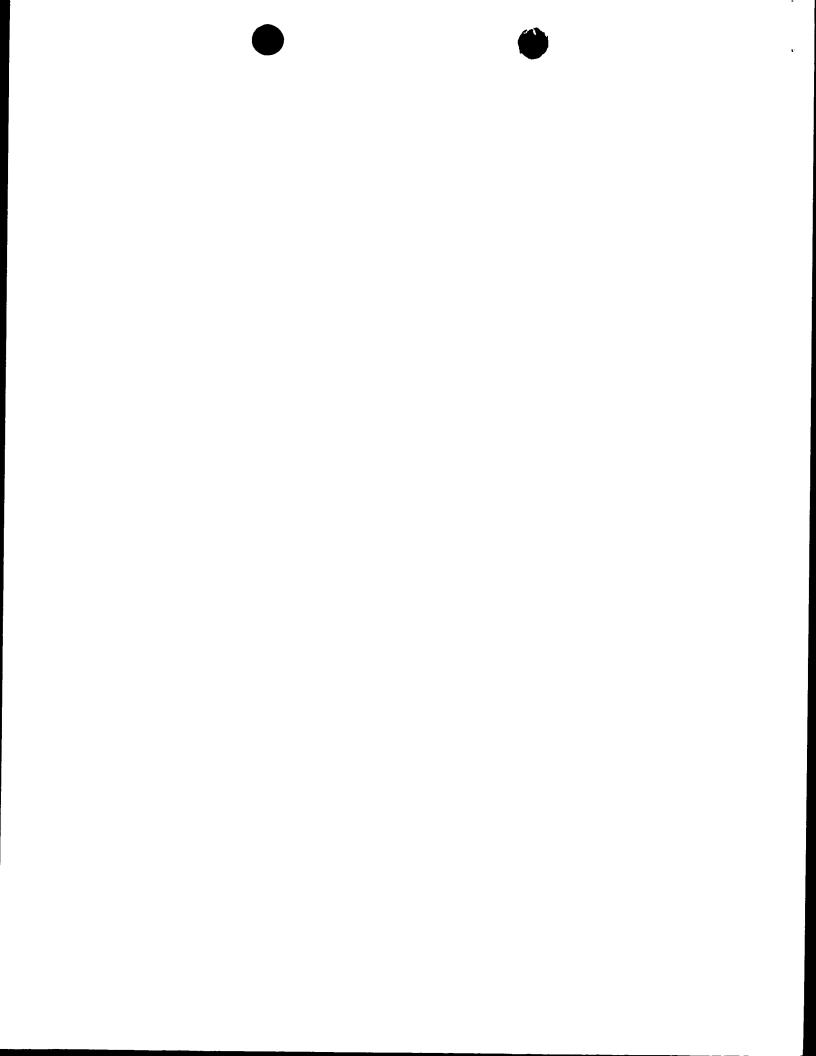
#### Neue Ansprüche 1-20:

- 1. Ein- oder mehrzelliger Mikroorganismus zur biotechnischen Herstellung von Riboflavin, dadurch gekennzeichnet, daß er eine Aktivität eines NAD(P)H-bildenden. Enzyms aufweist, die höher ist als diejenige eines Wildtyps der Species Ashbya gossypii ATCC10895.
- Ein- oder mehrzelliger Mikroorganismus nach Anspruch 1,
   dadurch gekennzeichnet, daß er eine erhöhte Isocitrat-Dehydrogenase-Aktivität aufweist.
- 3. Ein- oder mehrzelliger Mikroorganismus nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß er ein Pilz ist.
- 4. Ein- oder mehrzelliger Mikroorganismus nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß er ein Pilz aus der Gattung Ashbya ist.
- Ein- oder mehrzelliger Mikroorganismus nach einem der Ansprüche 1 bis 4,
   dadurch gekennzeichnet, daß er ein Pilz der Spezies Ashbya gossypii ist.
- 6. Isocitrat-Dehydrogenase-Gen mit einer für die in Fig. 11 (SEQ ID No. 2) angegebenen Aminosäuresequenz und deren Allelvariationen kodierenden Nukleotidsequenz.
- 7. Isocitrat-Dehydrogenase-Gen nach Anspruch 6 mit der Nukleotidsequenz von Nukleotid 1 bis 1262 gem. der Fig. 11 (SEQ ID No. 1).
- 8. Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 oder 7 mit einem vorgeschalteten Promotor der Nukleotidsequenz mit Nukleotid –661 bis –1 gem. der Fig. 11 (SEQ ID No. 1).
- Gen-Struktur enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche
   6 bis 8 sowie mit diesem operativ verknüpfte regulatorische Sequenzen.
- Vektor enthaltend ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 8 oder eine Gen-Struktur nach Anspruch 9.

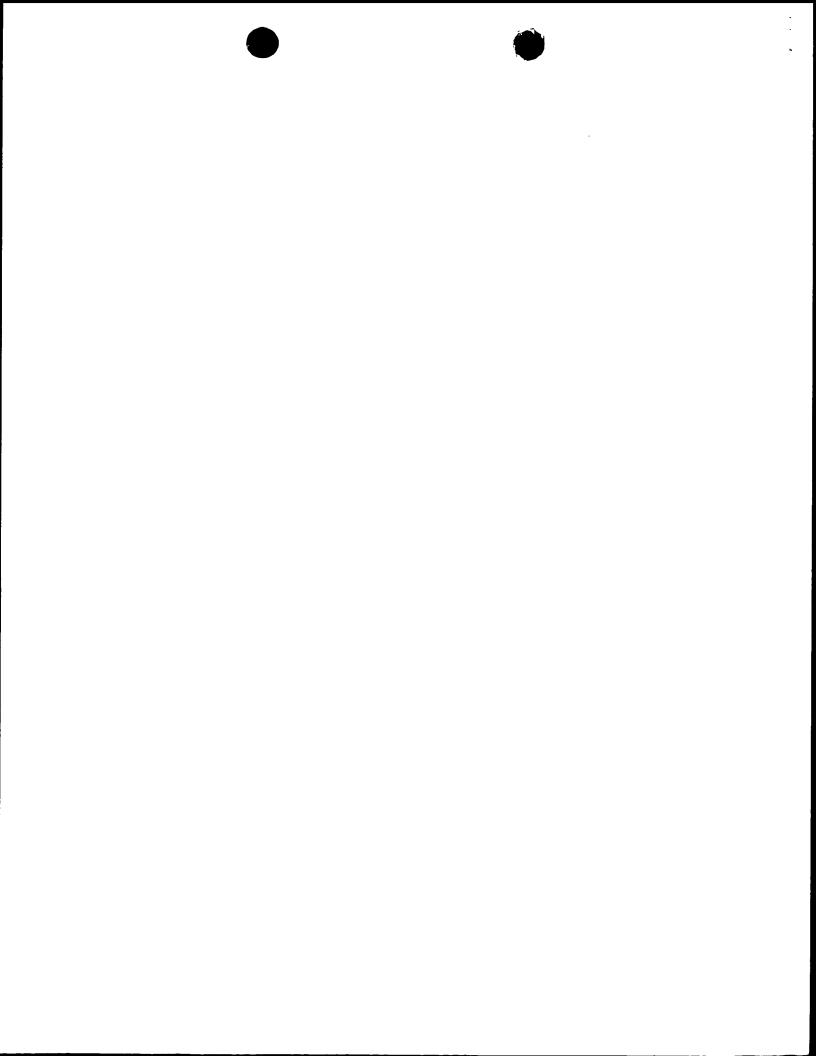




- 11. Genetisch veränderter Mikroorganismus zur biotechnischen Herstellung von Riboflavin enthaltend in replizierbarer Form ein Isocitrat-Dehydrogenase-Gen nach einem der Ansprüche 6 bis 8, welches im Vergleich zu dem entsprechend nicht genetisch veränderten Mikroorganismus verstärkt exprimiert wird und/oder deren Kopienzahl erhöht ist.
- 12. Genetisch veränderter Mikroorganismus gemäß Anspruch 11 enthaltend eine Gen-Struktur nach Anspruch 9 oder einen Vektor nach Anspruch 10.
- 13. Genetisch veränderter Mikroorganismus nach einem der Ansprüche 11 oder 12 enthaltend eine Isocitrat-Dehydrogenase, die im Vergleich zu dem entsprechend nicht genetisch veränderten Mikroorganismus eine gesteigerte katalytische Aktivität und/oder eine verringerte Hemmbarkeit gegenüber Inhibitoren aufweist.
- 14. Verfahren zur biotechnischen Herstellung von Riboflavin, dadurch gekennzeichnet, daß ein Mikroorganismus gemäß einem der Ansprüche 1 bis 5 oder 11 bis 13 eingesetzt wird.
- 15. Verfahren zur Herstellung eines Riboflavin produzierenden ein- oder mehrzelligen Mikroorganismus, dadurch gekennzeichnet, daß durch gentechnische Methoden die Aktivität eines NAD(P)H-bildenden Enzyms im Vergleich zu derjenigen eines Wildtyps der Species Ashbya gossypii ATCC10895 erhöht wird.
- 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß die Erhöhung der Enzymaktivität durch Austausch des Promotors und/oder Erhöhung der Genkopienzahl erzielt wird.
- 17. Verfahren nach einem der Ansprüche 15 oder 16, dadurch gekennzeichnet, daß die Erhöhung der Enzymaktivität durch eine gesteigerte katalytische Aktivität und/oder eine verringerte Hemmbarkeit gegenüber Inhibitoren der Isocitrat-Dehydrogenase erzeugt wird.
- 18. Verwendung eines Mikroorganismus nach einem der Ansprüche 1 bis 5 sowie 11 bis 13 zur Herstellung von Riboflavin.



- 19. Verwendung eines Isocitrat-Dehydrogenase-Gens nach einem der Ansprüche 6 bis 8 zur Herstellung eines Mikroorganismus nach einem der Ansprüche 1 bis 5 sowie 11 bis 13.
- Verwendung einer Gen-Struktur nach Anspruch 9 oder eines Vektors nach Anspruch
   10 zur Herstellung eines Mikroorganismus nach einem der Ansprüche 1 bis 5 sowie
   11 bis 13.

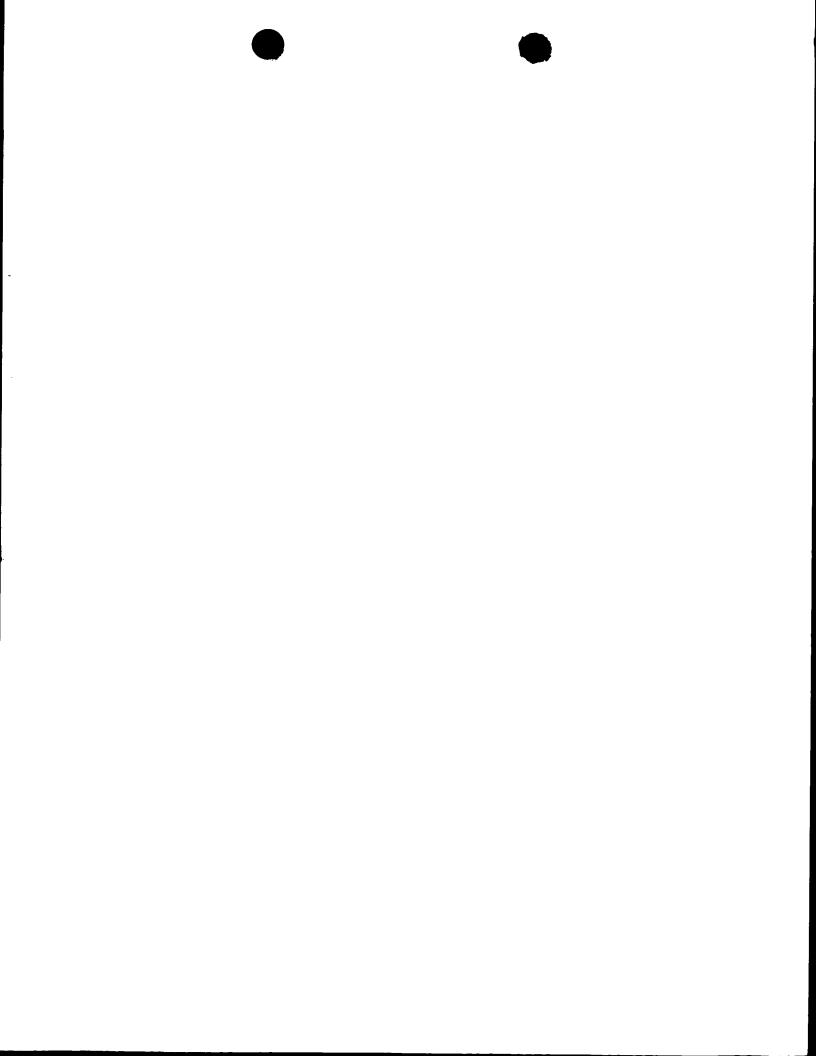


# **PCT**

# INTERNATIONALER RECHERCHENBERICHT

(Artikel 18 sowie Regeln 43 und 44 PCT)

Aktenzeichen des Anmelders oder Anwalts	WEITERES SIE	ne Mitteilung über die	Ubermittlung des internationalen rmblatt PCT/ISA/220) sowie, soweit
FZJ 9909 PCT	VORGEHEN zut	reffend, nachstehend	er Punkt 5
Internationales Aktenzeichen	Internationales Anmeldeda	tum (	(Frühestes) Prioritätsdatum (Tag/Monat/Jahr)
PCT/EP 00/07370	(Tag/Monat/Jahr) 31/07/2000	)	09/08/1999
Anmelder			
BASF AKTIENGESELLSCHAFT et	a1.		
Dieser internationale Recherchenbericht wurd Artikel 18 übermittelt. Eine Kopie wird dem Int	de von der Internationalen Re ternationalen Büro übermittel	cherchenbehörde ers t.	stellt und wird dem Anmelder gemäß
Dieser internationale Recherchenbericht umfa  X Darüber hinaus liegt ihm jev	aßt insgesamt <u>3</u> veils eine Kopie der in diesen	Blätter. n Bericht genannten U	Unterlagen zum Stand der Technik bei.
Grundlage des Berichts			
<ul> <li>a. Hinsichtlich der Sprache ist die inte durchgeführt worden, in der sie eing</li> </ul>	rnationale Recherche auf der gereicht wurde, sofern unter d	Grundlage der intern liesem Punkt nichts a	nationalen Anmeldung in der Sprache Inderes angegeben ist.
Die internationale Recherch Anmeldung (Regel 23.1 b))	e ist auf der Grundlage einer durchgeführt worden.	bei der Behörde eing	gereichten Übersetzung der internationalen
Recherche auf der Grundlage des S	Sequenzprotokolls durchgefül	art worden, das	Aminosāuresequenz ist die internationale
	Idung in Schrifticher Form en		orojoht worden ist
l <del>L</del>	onalen Anmeldung in comput		ereicht worden ist.
l ——	ch in schriftlicher Form einger ch in computerlesbarer Form (		<del>t</del>
Die Erklärung, daß das nac	hträglich eingereichte schriftli	che Sequenzprotokol	ll nicht über den Offenbarungsgehalt der
The state of the s	im Anmeldezeitpunkt hinausç omputerlesbarer Form erfaßte		schriftlichen Sequenzprotokoll entsprechen,
wurde vorgelegt.	·		
2. Bestimmte Ansprüche ha	ben sich als nicht recherch	<b>ierbar erwiesen</b> (siel	he Feld I).
3. Mangelnde Einheitlichkei	t der Erfindung (siehe Feld I	I).	
4. Hinsichtlich der Bezeichnung der Erfir	ndung		
X wird der vom Anmelder ein	gereichte Wortlaut genehmigt	· <b>.</b>	
wurde der Wortlaut von der	Behörde wie folgt festgesetz	t:	
•			
5. Hinsichtlich der Zusammenfassung	aaraiahta Wartlaut genehmig	•	
wurde der Wortlaut nach B	e innerhalb eines Monats nac	ngegebenen Fassung	g von der Behörde festgesetzt. Der sendung dieses internationalen
6. Folgende Abbildung der <b>Zeichnungen</b>	ist mit der Zusammenfassun	g zu veröffentlichen: A	
wie vom Anmelder vorgesc			X keine der Abb.
weil der Anmelder selbst ke	eine Abbildung vorgeschlager	n hat.	
weil diese Abbildung die Er	findung besser kennzeichnet		



#### INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen PEP 00/07370

A. KLASSIFIZIERUNG DES ANMELDUNGS GENSTANDES IPK 7 C12N15/53 C12N15/80 C12N9/04 //(C12N1/15,C12R1:645)

C12N1/15

C12P25/00

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

#### **B. RECHERCHIERTE GEBIETE**

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole ) IPK - 7 - C12N - C12P

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

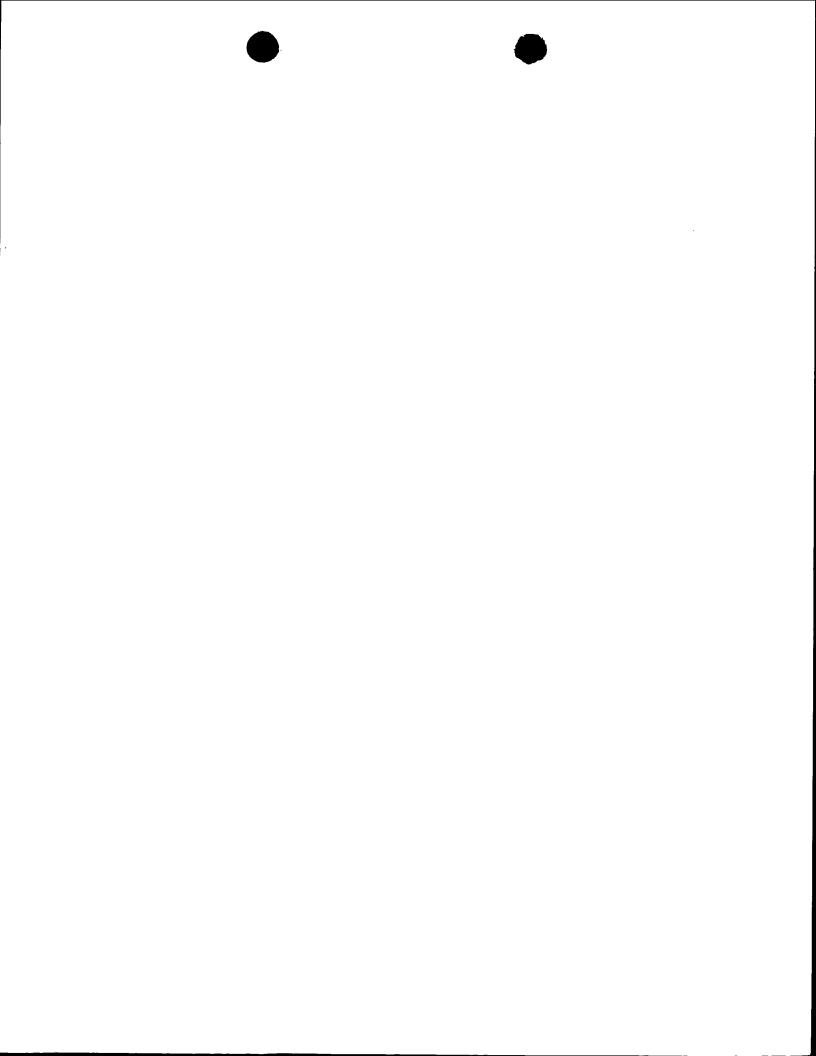
Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

BIOSIS, EPO-Internal, MEDLINE

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	HENKE BIRGIT ET AL: "IDP3 encodes a peroxisomal NADP-dependent isocitrate dehydrogenase required for the beta-oxidation of unsaturated fatty acids."  JOURNAL OF BIOLOGICAL CHEMISTRY, Bd. 273, Nr. 6, 6. Februar 1998 (1998-02-06), Seiten 3702-3711, XP002157778 ISSN: 0021-9258 Seite 3705; Abbildung 3 Seite 3709, Spalte 2 -Seite 3710  -/	7-10

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen	X Siehe Anhang Patentfamilie
<ul> <li>Besondere Kategorien von angegebenen Veröffentlichungen:</li> <li>"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist</li> <li>"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist</li> <li>"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)</li> <li>"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht</li> <li>"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist</li> </ul>	<ul> <li>*T* Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist</li> <li>*X* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden</li> <li>*Y* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist</li> <li>*&amp;* Veröffentlichung, die Mitglied derselben Patentfamilie ist</li> </ul>
Datum des Abschlusses der internationalen Recherche	Absendedatum des internationalen Recherchenberichts
19. Januar 2001	01/02/2001
Name und Postanschrift der Internationalen Recherchenbehörde	Bevollmächtigter Bediensteter
Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Blanco Urgoiti, B

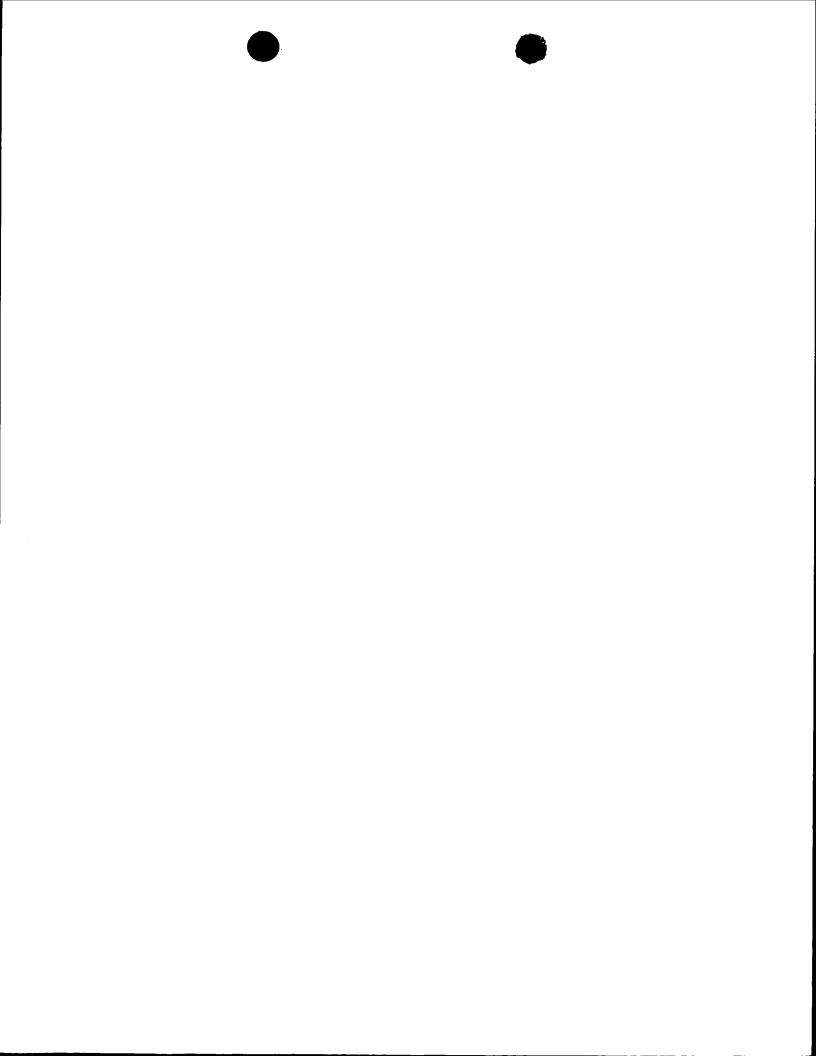
2



### INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzelchen
PEP 00/07370

	to the second se	Botr Angerrich Nr.
Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	WO 97 03208 A (BASF AG;KERNFORSCHUNGSANLAGE JUELICH (DE); KAESLER BRUNO (DE); SA) 30. Januar 1997 (1997-01-30) in der Anmeldung erwähnt das ganze Dokument	
A	EP 0 927 761 A (BASF AG) 7. Juli 1999 (1999-07-07) das ganze Dokument	
A	EP 0 405 370 A (HOFFMANN LA ROCHE) 2. Januar 1991 (1991-01-02) das ganze Dokument	

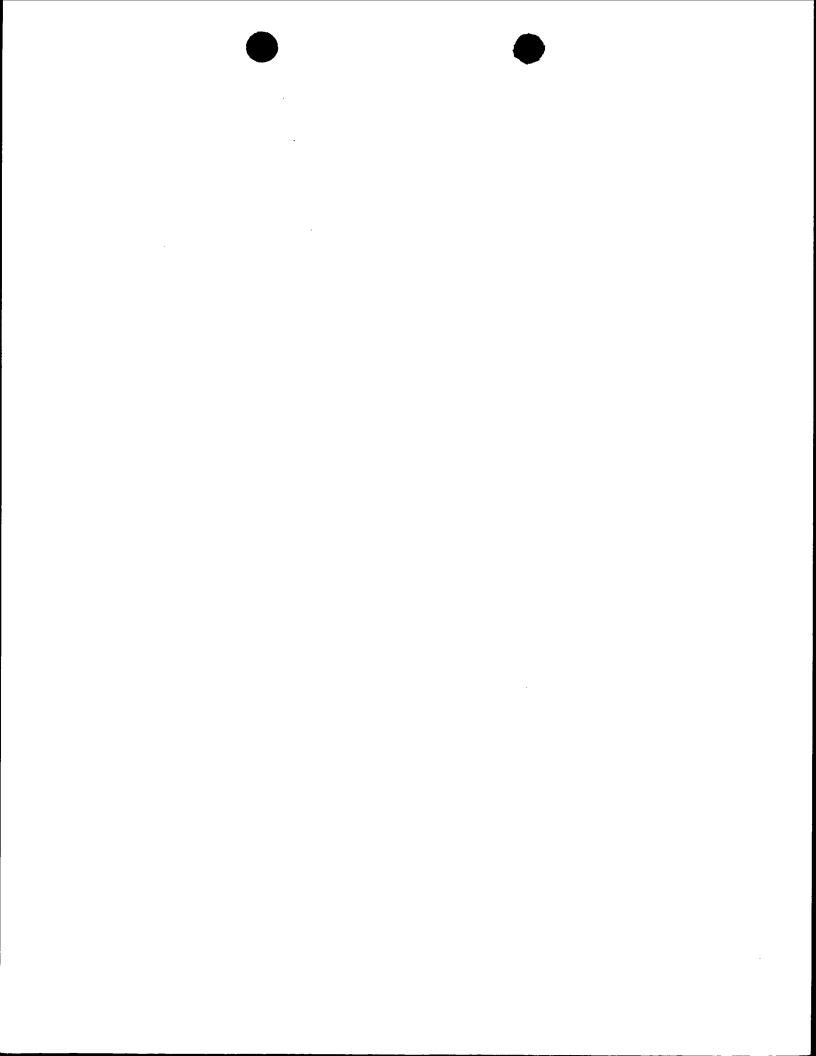


#### INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No				
PEP	00/07370			

Patent document cited in search report		Publication date	Patent family member(s)		Publication date	
WO	9703208	A	30-01-1997	DE DE CA CN EP JP US	19525281 C 19545468 A 2223877 A 1193356 A 0839211 A 11509409 T 5976844 A	04-04-1996 21-08-1997 30-01-1997 16-09-1998 06-05-1998 24-08-1999 02-11-1999
EP	0927761	Α	07-07-1999	CN WO EP JP	1227870 A 9933993 A 1040193 A 11243975 A	08-09-1999 08-07-1999 04-10-2000 14-09-1999
EP	0405370	A	02-01-1991	AT CN DE EP JP JP US US	195971 T 1049185 A 69033616 D 1001026 A 3117489 A 10066562 A 5925538 A 5837528 A	15-09-2000 13-02-1991 05-10-2000 17-05-2000 20-05-1991 10-03-1998 20-07-1999 17-11-1998



© 1998 by The American Society for Biochemistry and



3, No. 6, Issue of February 6, pp. 3702–3711, 1998 Printed in U.S.A.

# IDP3 Encodes a Peroxisomal NADP-dependent Isocitrate Dehydrogenase Required for the $\beta$ -Oxidation of **Unsaturated Fatty Acids\***

(Received for publication, September 5, 1997)

# Birgit Henke, Wolfgang Girzalsky, Veronique Berteaux-Lecellier‡, and Ralf Erdmann§

From the Department of Physiological Chemistry, Ruhr-Universität Bochum, 44780 Bochum, Germany

In Saccharomyces cerevisiae the metabolic degradation of saturated fatty acids is exclusively confined to peroxisomes. In addition to a functional  $\beta$ -oxidation system, the degradation of unsaturated fatty acids requires auxiliary enzymes, including a Δ2,Δ3-enoyl-CoA isomerase and an NADPH-dependent 2,4-dienoyl-CoA reductase. We found both enzymes to be present in yeast peroxisomes. The impermeability of the peroxisomal membrane for pyrimidine nucleotides led to the question of how the NADPH needed by the reductase is regenerated in the peroxisomal lumen. We report the identification and functional analysis of the IDP3 gene product, which is a yeast peroxisomal NADP-dependent isocitrate dehydrogenase. The newly identified peroxisomal protein is homologous to the mitochondrial Idp1p and cytosolic Idp2p, which both are yeast NADPdependent isocitrate dehydrogenases. Yeast cells lacking Idp3p grow normally on saturated fatty acids, but growth is impaired on unsaturated fatty acids, indicating that the peroxisomal Idp3p is involved in their metabolic utilization. The data presented are consistent with the assumption that peroxisomes of S. cerevisiae contain the enzyme equipment needed for the degradation of unsaturated fatty acids, including an NADPdependent isocitrate dehydrogenase, a putative constituent of a peroxisomal NADPH-regenerating redox system.

Peroxisomes harbor variable metabolic pathways that differ depending on cell type, developmental stage, and food supply (1, 2). In reference to the multiplicity of cellular functions and to the ability of cells to adjust the enzymatic equipment as well as the size and number of these organelles in response to the cellular demand, peroxisomes are appropriately called multipurpose organelles (3). The importance of peroxisomes for cellular function is especially emphasized by a number of inherited diseases in humans that are caused by peroxisomal dysfunction and usually have profound clinical consequences (4).

A typical metabolic pathway of peroxisomes is the  $\beta$ -oxidation of fatty acids (5, 6). In fact, whereas the presence of a mitochondrial β-oxidation system is restricted to mammalian

cells and a few protists (7), the fatty acid oxidation in peroxisomes is nearly ubiquitous among eukaryotic cells (7, 8). The peroxisomal and the mitochondrial degradation of fatty acids is performed by functionally comparable but genetically distinct proteins (8, 9). In fungi and plants, the degradation of fatty acids exclusively takes place in peroxisomes, and growth on fatty acids results in proliferation of peroxisomes accompanied by a massive induction of peroxisomal proteins including the  $\beta$ -oxidation enzymes (7, 10).

In addition to the chain shortening  $\beta$ -oxidation system, the oxidation of unsaturated fatty acids requires auxiliary enzymes for the elimination of the double bonds (8). Degradation of unsaturated fatty acids with odd-numbered double bonds requires a  $\Delta 2, \Delta 3$ -enoyl-CoA isomerase (Fig. 1B) (11). For the degradation of unsaturated fatty acid with even-numbered double bonds, an NADPH-dependent 2,4-dienoyl-CoA reductase is needed in addition to the isomerase (Fig. 1A) (12). Recently, a novel pathway for the degradation of unsaturated fatty acids with double bonds at odd-numbered carbon atoms has been described that also requires the NADPH-dependent reductase described above (Fig. 1C) (13, 14). The successive reduction and isomerization of double bonds by these auxiliary enzymes results in the formation of intermediates that can reenter the  $\beta$ -oxidation spiral (8). The presence of both the Δ2,Δ3-enoyl-CoA isomerase and the NADPH-dependent 2,4dienoyl-CoA reductase has been demonstrated in all peroxisomes studied so far (8). As the peroxisomal membrane has been suggested to be impermeable for small solutes (15), the requirement of the peroxisomal enoyl-CoA reductase for NADPH raises the question of the existence of an NADPH regenerating system in peroxisomes.

We applied a reverse genetic approach to identify proteins essential for peroxisome function in Saccharomyces cerevisiae. Here we report the identification, characterization, and functional analysis of a peroxisomal NADP-dependent isocitrate dehydrogenase. Deficiency in this enzyme resulted in an impaired growth of S. cerevisiae on unsaturated fatty acids, whereas growth on saturated fatty acids was not affected. The peroxisomal Idp3p is suggested to be involved in the regeneration of the NADPH needed for the peroxisomal degradation of unsaturated fatty acids.

#### EXPERIMENTAL PROCEDURES

Strains, Growth Conditions, and General Methods-The yeast strains used in this study were S. cerevisiae wild-type UTL-7A (MATa, ura3-52, trp1, leu2-3,112), pex7Δ (MATa, pex7::LEU2, ura3-52, trp1)) (16), idp3\(Delta\) (MATa, ura3-52, trp, leu2-3,112, idp3::kanMX4) (this paper), idp1\(\Delta\) (MATa, leu2-3,112, his3-1, trp1-289, idp1::URA3) (17), idp1,idp3Δ (MATa, leu2-3,112, his3-1, trp1-289, idp1::URA3, idp3::hanMX4) (this paper), idp1/idp2\( (MATa, leu2-3,112, his3-1, trp1-289, idp1::URA3, idp2::URA3) (18), and idp1/idp2/idp3Δ (MATa, leu2-3,112, his3-1, trp1-289, idp1::URA3, idp2::URA3, idp3::kanMX4)

Yeast strains were grown at 30 °C in YPD or in minimal medium

‡ Present address: Institut de Genetique et Microbiologie, Universite

Paris Sud, 91405 Orsay Cedex, France.

<sup>\*</sup> This work was supported by Deutsche Forschungsgemeinschaft Grant Er 178/2-1. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>§</sup> To whom correspondence should be addressed: Ruhr-Universität Bochum, Insitut für Physiologische Chemie, Universitätsstr. 150, 44780 Bochum, Germany. Tel.: 0234-700-4947; Fax: 0234-709-4279; E-mail: Ralf. Erdmann@rz.ruhr-uni-bochum.de.

Fig. 1. Pathways of the peroxisomal degradation of unsaturated fatty acids. In addition to the  $\beta$ -oxidation spiral, auxiliary enzymes are required to degrade unsaturated fatty acids. A, fatty acids with double bonds at even-numbered carbon atoms require an NADP-dependent 2,4-dienoyl reductase and  $\Delta 2,\Delta 3$ -enoyl-CoA isomerase (8). B, degradation of fatty acids with double bonds extending from odd-numbered carbon atoms at least requires the activity of the  $\Delta 2,\Delta 3$ -enoyl-CoA isomerase (11). C, in the alternative pathway discovered by Smeland et al. (13), the NADP-dependent 2,4-dienoyl reductase is also needed for the degradation of fatty acids with odd-numbered double bonds.

**B-oxidation** 

(SD) as described previously (19). For oleic acid induction of peroxisome proliferation, cells were grown in SD medium to late log phase and then shifted into YNO (20) and incubated for 14 h. Necessary auxotrophic requirements were added according to Ausubel et al. (21).

Whole yeast cell extracts were prepared by the method of Yaffe and Schatz (22). Enzymatic modification of DNA, fragment purification, and bacterial transformation were performed essentially as described by Ausubel et al. (21). Yeast transformations were carried out according to Gietz and Sugino (23).

Purification and Amino Acid Sequencing of Idp3p-High salt-extracted peroxisomal membranes were prepared from oleic acid-induced SKQ2N cells. Further separation of the peroxisomal membrane proteins was achieved by reverse-phase HPLC1 according to Erdmann and Blobel (24)

For sequencing of Idp3p, the SDS samples of HPLC fractions containing the protein (fractions 42-46; see Fig. 2) were pooled and separated on a 12% SDS-polyacrylamide gel. Polypeptides were electrophoretically transferred onto a polyvinylidene difluoride membrane and visualized with 0.1% Amido Black in 10% acetic acid. Idp3p was excised, and Lys-C-derived peptides of the protein were separated by HPLC and subjected to sequence analysis on a gas phase sequenator according to Fernandez et al. (25). Protein sequence analysis was provided by the Rockefeller University Protein Sequencing Facility, which is supported in part by National Institutes of Health shared instrumentation grants and by funds provided by the U.S. Army and Navy for the purchase of equipment.

Isolation and Sequencing of IDP3-According to the obtained internal sequences of Idp3p, degenerated sense (5'-GCGAATTCA(C/T)C-CIAT(A/C/T)GT(A/G/C/T)GA(A/G)ATG-3') and antisense (5'-TCTAAG- CTT(A/G/C/T)GCIAC(C/T)TC(A/G)TC(T/G/A)AT-3') oligonucleotide primers, distinguishing between the IDP3 gene and the highly homologous IDP1 and IDP2 genes were designed. The corresponding genomic region of the IDP3 gene was amplified by the polymerase chain reaction with yeast genomic DNA (100 ng; Promega, Madison, WI) as template. The amplification product was isolated and subcloned into a derivative pBluescript SK(+) using the Srf1 kit (Stratagene, La Jolla, CA), resulting in pSRF-IDP3. The authenticity of the insert was confirmed by DNA sequencing. A [32P]dATP-labeled probe of 520 base pairs was generated by PCR with oligonucleotide primers iso6 (5'-GCCACTATAACAC-CCGATG-3') and iso1 (5'-CGTACGTTATTTTTAAAGCCTG-3') and the plasmid pSRF-IDP3 as template, using a random-primed labeling kit (Boehringer, Mannheim, Germany). High stringency hybridization according to Maniatis et al. (26) was performed to screen a YEp13-based yeast genomic library (27) that was kindly provided by M. Bolotin-Fukuhara. Two positive clones containing the IDP3 gene were isolated, and sequencing was directly performed on one of the plasmids, YEp13-IDP3, with an automatic sequencer (model 373A; Applied Biosystems, Weiterstadt, Germany), the DyeDesoxy terminator cycle sequencing kit (Applied Biosystems), and synthetic oligonucleotides. Both strands of the IDP3 gene were sequenced.

Disruption of the Genomic IDP3 Gene-To construct a idp3 null mutant, the entire IDP3 open reading frame was replaced by the kanMX4 marker gene (28). The PCR-derived construct for disruption comprised the kanMX4 gene flanked by short (40-base pair) homology regions to the IDP3 3'- and 5'-noncoding region. PCR primers were 5'-CACAAGCAACACTTTAGAGATAGTTGTCCAAGTTAAAATGCG-TACGCTGCAGGTCGAC-3' (KU179) and 5'-GGCCAGACTTGTCTTT-TCAAATGAATGGCGGATTGGTTTAATCGATGAATTCGAGCTCG-3' (KU180), and plasmid pFA6a-kanMX4 served as template. The resulting amplification construct was introduced into S. cerevisiae wild-type UTL-7A, the  $idp1\Delta$  mutant (17), and the  $idp1\Delta/idp2\Delta$  double mutant (18). Geneticin-resistant clones were selected by growth on YPD plates containing 200 mg/liter G418 (28).

7

<sup>1</sup> The abbreviations used are: HPLC, high pressure liquid chromatography; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; PMSF, phenylmethylsulfonyl fluoride; PTS1, peroxisomal targeting signal 1.

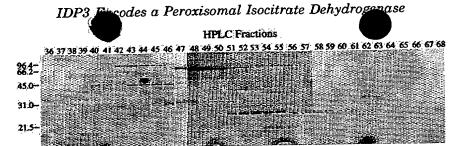


Fig. 2. Preparative chromatographic separation of peroxisomal membrane proteins. High salt-extracted peroxisomal membranes (1 mg of protein) were solubilized in SDS and separated by reverse phase HPLC. Polypeptides of selected fractions were separated by SDS-PAGE and visualized by Coomassie Blue staining. The band of Idp3p is indicated by the arrowhead. The amount per lane corresponded to 5% of the total fraction. Molecular mass standards are indicated on the left.

Antibodies and Immunoblots-For generating antibodies against Idp3p, the IDP3-fragment IDP3\* (144-210 amino acids) was amplified by PCR using GenBank™ plasmid YEp13-IDP3 and oligonucleotides KU188 (5'-CCGGAATTCCGGGATCCGATCAAGATTAAAAAAGCA-3') and KU189 (5'-TGCTCTAGACTGCAGCTACTCGAGTGTAAAGAAT-AACGGTAG-3'). The PCR product was digested with EcoRI and XbaI and inserted into pBluescript SK(+) (Stratagene, La Jolla, CA) to create pSKIDP3\*. The BamHI/HindIII fragment of pSKIDP3\* was subcloned into pET21b (Novagen), leading to plasmid pET-IDP3\*. The plasmid was transformed into Escherichia coli BL21-DE3, resulting in an isopropyl-1-thio-β-D-galactopyranoside-inducible expression of HIS<sub>e</sub>tagged Idp3p\*. The tagged Idp3p\* was purified by affinity chromatography on a nickel-nitrilotriacetic acid resin (Quiagen, Hilden, Germany) according to the manufacturer's protocol. Rabbit polyclonal antibodies to HIS<sub>6</sub>-tagged Idp3p\* were produced by Eurogentec (Seraing, Belgium) according to standard methods (29).

Anti-thiolase (Fox3p), anti-Kar2p, anti-Pcs60p, anti-Pex14p, and anti-Pex3p antibodies have been described previously (30-34). Anti-rabbit or anti-mouse IgG-coupled horseradish peroxidase (Amersham Corp.) were used as the second antibody, and blots were developed using the ECL system (Amersham). Western blot analyses were performed ac-

cording to standard protocols (29).

Construction and Expression of Idp3pΔCKL—To construct a Idp3pΔCKL, the IDP3 gene was amplified by PCR using plasmid YEp13-IDP3 (see above) and oligonucleotides KU308 5'-CCGCTCGA-GGGCTGGTGAAAAGACAGT-3' and KU311 5'-CGCGGATCCTTAC-ATACCTTTCTTGTCTTCAT-3'. The amplification product was isolated and subcloned (BamHIXhoI) into the vector pRSterm, a pRS316 derivative (35) that contained the HincII/KpnI-prepared CYCI terminator of pTerm1 (36). The resulting plasmid pRS-IDP3ΔCKL was transformed into idp3Δ, resulting in an expression of Idp3pΔCKL under the control of its own promoter.

Cell Fractionation—Spheroplasting of yeast cells, homogenization, and differential centrifugation at  $25,000 \times g$  of homogenates were

performed as described previously (19).

For subfractionation by isopycnic sucrose density gradient centrifugation, cell lysates or organellar pellets were loaded onto linear 20-53% sucrose density gradients (34). Centrifugation, fractionation of gradients, and preparation of the samples for SDS-PAGE were carried out as described (24).

The suborganellar localization of proteins was determined by extraction of 25,000  $\times$  g organellar pellets with low salt (10 mm Tris/HCl, pH 8.0; 1 mm PMSF), high salt (10 mm Tris/HCl, pH 8.0; 500 mm KCl; 1 mm PMSF), or pH 11 buffer (0.1 m Na<sub>2</sub>CO<sub>3</sub>, 1 mm PMSF) as described by

Erdmann and Blobel (37).

Protease Protection—Peroxisomal peak fractions from a sucrose density gradient were pooled and diluted five times in gradient buffer (34). Peroxisomes were sedimented at  $25,000 \times g$  for 30 min and subsequently resuspended in homogenization buffer (19) without protease inhibitors but supplemented with 50 mM KCl. Equal amounts of isolated peroxisomes were incubated with increasing amounts of proteinase K for 10 min on ice. Protease was inactivated immediately after the incubation with 4 mM PMSF, proteins were precipitated with trichloroacetic acid, and samples were processed for SDS-PAGE.

Heterologous Expression, Purification, and Characterization of Idp3p—The IDP3-orf was amplified by PCR using plasmid YEp13-IDP3 as template and oligonucleotides KU235 (5'-CGGAATTCCCATATGAGTAAAATTAAAGTTGTT-3') and KU236 (5'-CCCTCGAGTAGTTTGCA-CATACCTTTC-3'). The PCR product was digested with EcoRI and XhoI and subcloned into pET21b (Novagen), leading to plasmid pET-IDP3. The plasmid was introduced into E. coli BL21-DE3, resulting in an isopropyl-1-thio-β-D-galactopyranoside-inducible expression of HIS<sub>g</sub>-

tagged Idp3p. The tagged Idp3p was purified by affinity chromatography on a nickel-nitrilotriacetic acid resin (Quiagen, Hilden, Germany) according to the manufacturer's protocol.

Analytical Procedures—NADP-specific isocitrate dehydrogenase activity was determined as described by Loftus et al. (18). Catalase and fumarase were assayed as described by Moreno de la Garca et al. (38). Δ3,Δ2-Enoyl-CoA isomerase was assayed spectrophotometrically at 340 nm in a coupled assay (39) with 3-trans-decanoyl-CoA as substrate. 2,4-Dienoyl-CoA reductase activity was determined spectrophotometrically at 340 nm with 2-trans-decanoyl-CoA as substrate according to Kunau and Dommes (12). Total protein was measured by the BCA method (Pierce) using bovine serum albumin as a standard.

#### RESULTS

Isolation and Identification of Idp3p-Peroxisomes were isolated from oleic acid-induced S. cerevisiae cells and successively extracted by low and high salt treatments. The proteins of high salt extracted peroxisomes were solubilized with SDS and separated by HPLC and SDS-PAGE (Fig. 2). Lys-C-derived internal fragments of the 45-kDa protein marked in Fig. 2 were subjected to amino-terminal protein sequencing in preparation for DNA cloning and sequencing of the corresponding gene (see "Experimental Procedures"). The open reading frame of the isolated DNA fragment encoded a new protein with a calculated molecular mass of 48 kDa (Fig. 3) that later on also appeared in the yeast genome data base as open reading frame YNL009w. A search of the GenBank<sup>TM</sup> data base with the predicted amino acid sequence of IDP3 revealed the yeast genes IDP1 (17) and IDP2 (18) as close relatives of the newly identified gene, hereafter referred to as IDP3. The overall identity of Idp3p with Idp1p and Idp2p is 68 and 70%, respectively (Fig. 4). Idp1p and Idp2p represent the two NADP-dependent isocitrate dehydrogenases reported for S. cerevisiae to date. Idp2p is localized in the yeast cytosol, and Idp1p is a mitochondrial isoenzyme that differs from the other two proteins by an N-terminal extension, which functions as a mitochondrial targeting signal (Fig. 4) (18, 40). Idp3p lacks the mitochondrial targeting signal and instead is characterized by an additional nine amino acids at the extreme C terminus. These terminal amino acids of Idp3p comprise the tripeptide cysteine-lysineleucine (CKL; Fig. 3), a putative peroxisomal targeting signal 1 (PTS1) for S. cerevisiae (41, 42). The prominent presence in the HPLC profile of peroxisomal proteins (Fig. 2), the sequence similarity to Idp1p and to Idp2p (Fig. 4), and the presence of the putative peroxisomal targeting signal (PTS1; Fig. 3) suggested that Idp3p might be a peroxisomal NADP-dependent isocitrate dehydrogenase.

IDP3 Is Induced upon Growth on Oleic Acid—Antibodies were generated against an internal fragment of Idp3p comprising amino acids 144–210, which displays the lowest similarity of the protein to Idp1p and Idp2p (Fig. 4). A polypeptide with the predicted molecular mass for Idp3p (48 kDa) was detected in wild-type but not in idp3Δ yeast extracts (Fig. 5A). Although binding of the antibodies to other proteins was observed under low stringency conditions, none of these bands disappeared in

Fig. 3. Nucleotide and deduced amino acid sequence of the S. cerevisiae IDP3 locus. A putative oleic acid response element (43) and a presumptive TATA sequence (54) in the 5'-noncoding region are underlined. The underlined amino acid sequences were obtained by peptide sequencing of protein fragments derived by Lys-C digestion of the isolated ldp3p. The putative PTS1 of the C terminus is double underlined.

mutants lacking either Idp1p or Idp2p, indicating that the antibodies generated do not recognize Idp1p or Idp2p but are specific for Idp3p (Fig. 5A).

In S. cerevisiae, growth on oleic acid results in a massive proliferation of peroxisomes accompanied by the induction of peroxisomal enzymes involved in peroxisomal fatty acid degradation (7, 10). The oleic acid induction is mediated by the transcription activator Pip2p, which binds to a well defined oleic acid-responsive element at the promoter of several peroxisomal proteins (43). A perfect consensus sequence for Pip2p binding is also present upstream of the IDP3 open reading frame (position -311 to -289; Fig. 3), and Idp3p was found to be highly inducible by oleic acid (Fig. 5B), supporting the assumption of Idp3p being involved in peroxisomal fatty acid degradation.

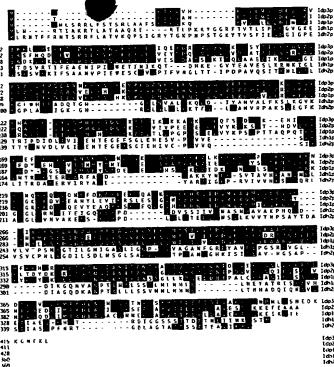


Fig. 4. Comparison of the deduced amino acid sequence of S. cerevisiae Idp3p with other isocitrate dehydrogenases from S. cerevisiae. The mitochondrial Idp1p (17) and the cytosolic Idp2p (18) are both NADP-dependent isocitrate dehydrogenases. The Idh1p (47) and Idh2p (48) are subunits of the heteromeric mitochondrial NAD-dependent isocitrate dehydrogenase. Alignment was performed with the BESTFIT program (EMBL, Heidelberg). Amino acids conserved in Idp3p and at least one of the other proteins are blocked.

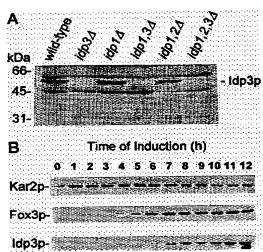


Fig. 5. Immunological detection of Idp3p and time course of Idp3p induction by oleic acid. A, equal amounts of whole-cell lysates from oleic acid-induced wild type and indicated single, double, or triple  $(idp1\Delta, idp2\Delta)$ , or  $idp3\Delta)$  mutant cells were subjected to Western blot analysis with rabbit antiserum against Idp3p. The amount loaded corresponds to 0.5% of extracts from 30 mg of cells. B, wild-type cells were precultured in 0.3% SD and subsequently shifted to oleic acid-containing medium. At the indicated time points, whole-cell extracts were prepared for immunological detection of oleic acid-inducible Fox3p (30), constitutively expressed Kar2p (31), and Idp3p. The amounts loaded correspond to 0.3 mg of cells.

Peroxisomes of S. cerevisiae Contain an NADP-dependent Isocitrate Dehydrogenase—The subcellular localization of NADP-dependent isocitrate dehydrogenases in S. cerevisiae

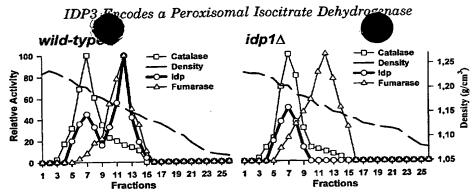


Fig. 6. Organellar localization of NADP-dependent isocitrate dehydrogenase isoenzymes in wild-type and  $idp1\Delta$  mutant cells. Organelles obtained by a  $25,000 \times g$  centrifugation of cell homogenates from oleic acid-induced wild-type and  $idp1\Delta$  mutant cells were separated on isopycnic 20-53% (w/w) sucrose density gradients. Fractions of 1 ml were collected from the bottom of the gradients. Relative amounts of the peroxisomal marker enzyme catalase and the mitochondrial fumarase as well as NADP-dependent isocitrate-dehydrogenase were monitored by activity measurements. Peroxisomes peaked in fraction 7 at a density of 1.21 g/ml, and mitochondrial peaked in fraction 12 at a density of 1.18 g/ml. NADP-dependent isocitrate-dehydrogenase activity was detected in both peroxisomal and mitochondrial fractions of the wild-type lysate but only in peroxisomal fractions of the  $idp1\Delta$  lysate.

was first analyzed by differential centrifugation of cell homogenates from oleic acid-induced wild-type yeast cells. More than 60% of the Idp activity was found in the supernatant fraction, suggesting that the cytosolic isoform might be responsible for the majority of the endogenous enzyme activity (data not shown). Sedimented organelles were further fractionated by sucrose density gradient centrifugation. NADP-dependent isocitrate dehydrogenase activity was found in both the mitochondrial and the peroxisomal fractions (Fig. 6). The peroxisomal peak, however, comprised a smaller fraction of the total particular enzyme activity. To exclude the possibility that the activity found in the peroxisomal fraction was due to a mitochondrial contamination, the subcellular localization of the enzymes was also analyzed in a  $idp1\Delta$  mutant strain, lacking the mitochondrial NADP-dependent isocitrate dehydrogenase. After differential centrifugation of cell homogenates of the  $idp1\Delta$  mutant strain, about 15% of the total NADP-dependent isocitrate dehydrogenase activity was still localized to the organellar fraction. Subsequent sucrose density gradient fractionation confirmed the absence of the mitochondrial enzyme and demonstrated the activity to exclusively co-segregate with peroxisomal marker enzymes (Fig. 6), consistent with the presence of a peroxisomal isoenzyme of the NADP-dependent isocitrate dehydrogenases in S. cerevisiae.

The Peroxisomal NADP-dependent Isocitrate Dehydrogenase Activity Is Performed by Idp3p-As a first step to analyze whether the peroxisomal NADP-dependent isocitrate activity is due to Idp3p, we studied the subcellular localization of the protein. Immunological detection of Idp3p in fractions generated by differential centrifugation of yeast cell homogenates revealed that the protein is exclusively found in the organellar pellet (Fig. 7A). Immunoblot analysis of fractions gained by subsequent sucrose density gradient centrifugation of the organellar pellets demonstrated the protein to be exclusively localized in the peroxisomal fractions (Fig. 7B). In agreement with Idp3p being responsible for the peroxisomal NADP-dependent isocitrate dehydrogenase activity, the absence of this protein in idp3A mutant cells correlated with the disappearance of the enzyme activity in the organellar pellet and in the peroxisomal fractions of sucrose density gradients (Fig. 7).

The sequence similarity of Idp3p to the two NADP-dependent isocitrate dehydrogenases suggested that the protein Idp3p itself is an NADP-dependent isocitrate dehydrogenase. However, to confirm that the lack of the peroxisomal NADP-dependent isocitrate dehydrogenase activity upon IDP3 deletion is not caused indirectly, the protein was heterologously expressed in E. coli, and the enzyme properties of isolated Idp3p were analyzed. Transformation of E. coli with a plasmid carrying the

IDP3 gene under the control of the bacterial promoter resulted in a massive increase in NADP-dependent isocitrate dehydrogenase activity in bacterial lysates accompanied by the appearance of immunoreactive Idp3p (Fig. 8A). Taking advantage of the C-terminal histidine tag, the protein was purified to apparent homogeneity as judged by SDS-PAGE. The isolated Idp3p showed a specific NADP-dependent isocitrate dehydrogenase activity of 1654 nanokatals/mg. Since expression of Idp3p did result in the concomitant appearance of NADP-dependent isocitrate dehydrogenase activity in bacterial extracts and since the enzyme activity was retained by purified Idp3p, these data confirmed Idp3p to be the yeast peroxisomal NADP-dependent isocitrate dehydrogenase (Fig. 8B).

A set of kinetic properties of the Idp3p enzyme activity were studied with the recombinant yeast protein purified from E. coli extracts. The enzyme activity did strongly depend on the presence of NADP<sup>+</sup> that could not be replaced by NAD<sup>+</sup> (data not shown). The  $K_m$  values for NADP<sup>+</sup> and isocitrate were 0.02 and 0.05 mm, respectively (Fig. 9), and are in the range of those reported for the peroxisomal NADP-dependent isocitrate dehydrogenase of Candida tropicalis (0.016 mm for NADP and 0.11 mm for isocitrate) (44).

Idp3p Is Localized in the Peroxisomal Lumen—An organellar fraction isolated from spheroplasts of yeast wild-type cells was subjected to extraction by low salt, high salt, and carbonate at pH 11 according to Ref. 37. Idp3p was resistant to low salt extraction but was released by high salt and carbonate treatment of the organelles (Fig. 10A). These extraction properties distinguished Idp3p from two other peroxisomal proteins. Pex3p was resistant to all treatments, consistent with it being an integral membrane protein (34). As expected for a matrix protein, peroxisomal thiolase (Fox3p) (30) was extracted by all treatments. The extractability of Idp3p by carbonate treatment suggested that Idp3p does not span the peroxisomal membrane. Idp3p also does not seem to be tightly associated with the peroxisomal membrane, since part of the protein could be extracted with high salt. The extraction properties of Idp3p are similar to those observed for Pcs60p, a protein of the peroxisomal matrix that is also loosely associated with the peroxisomal membrane (32). To distinguish whether Idp3p is associated with the outer aspects of peroxisomes or whether the protein resides in the peroxisomal lumen, we analyzed the sensitivity of organellar Idp3p to externally added proteases. In the presence of detergents and proteases, all proteins were rapidly degraded. When detergents were present, degradation of proteins was also observed without the addition of protease, presumably due to the liberation of endogenous proteases (data not shown). However, in the absence of detergents, both the

ſ

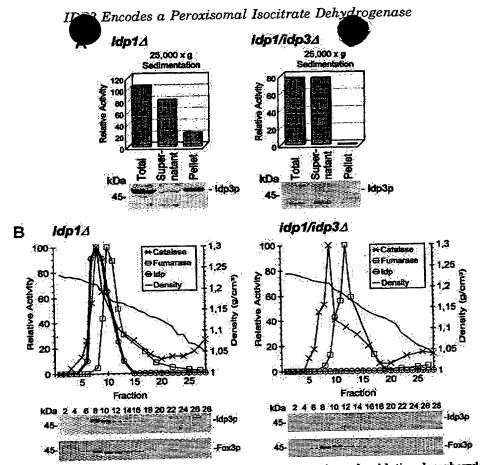


Fig. 7. Idp3p is localized in yeast peroxisomes and is required for the peroxisomal oxidative decarboxylation of isocitrate. A, immunoblot analysis and enzyme activity measurements of cell fractions that were obtained by differential centrifugation of cell-free extracts from oleic acid-induced  $idp1\Delta$  and  $idp1lidp3\Delta$  cells. Equal volumes of each fraction were immunologically analyzed for the presence of Idp3p. In parallel, the fractions were assayed for NADP-dependent isocitrate dehydrogenase activity. Idp3p was exclusively localized to the organellar fraction of  $idp1\Delta$  cells but was absent in  $idp1lidp3\Delta$  cells. In  $idp1\Delta$  cells, about 80% of the total enzyme activity was found in the soluble fraction. In  $idp1lidp3\Delta$  cells, the deficiency in Idp3p correlated with the disappearance of NADP-dependent isocitrate dehydrogenase activity in peroxisomal fractions obtained by isopycnic 20–53% (w/w) sucrose density gradient centrifugation of organelles of the 25,000 × g pellet from oleic acid-induced  $idp1\Delta$  and  $idp1lidp3\Delta$  cells. Equal volumes of each fraction were immunologically analyzed for the presence of Idp3p and thiolase. Relative amounts of NADP\*-dependent cells. Equal volumes of each fraction were immunologically analyzed for the presence of Idp3p and thiolase. Relative amounts of NADP\*-dependent enterior dehydrogenase and peroxisomal marker enzymes catalase and mitochondrial fumarase were monitored by enzyme activity measurements. Peroxisomes peaked at a density of 1.21 g/ml, and mitochondria peaked at a density of 1.17 g/ml. Idp3p was found predominantly in the peroxisomal peak fractions of  $idp1\Delta$  cells but was absent in  $idp1lidp3\Delta$  cells. No NADP-dependent isocitrate dehydrogenase activity was detected in peroxisomes lacking the Idp3p.

intraperoxisomal thiolase (Fox3p) and Idp3p were protected against added proteases (Fig. 10B). Under the same conditions, Pex14p, which is located at the cytosolic face of the peroxisomal membrane (33), was rapidly degraded. Taken together, these results are consistent with an intraperoxisomal localization of Idp3p.

Peroxisomal Targeting of Idp3p Depends on the Presence of the Three C-terminal Amino Acids—The amino acids CKL at the extreme C terminus of Idp3p fit the consensus for a yeast PTS1 (41, 42). To analyze whether this putative PTS1 of Idp3p is functional, we analyzed the subcellular localization of a mutated Idp3pΔCKL lacking the last three amino acids. Idp3p $\Delta$ CKL was expressed in an  $idp3\Delta$  strain, and localization of the protein was determined by subcellular fractionation of whole-cell homogenates on sucrose density gradients. Idp $3p\Delta$ CKL did not co-segregate with the peroxisomal markers but instead was exclusively found in the loading zone of the gradient, suggesting a cytosolic localization of the protein (data not shown). This result indicated that the last three amino acids of Idp3p are essential for the peroxisomal targeting of the protein. The presence of a functional PTS1 in Idp3p is in line with the observed protease resistance of the protein (Fig. 10B), since this signal sequence is known to target proteins to the peroxisomal matrix (41, 45).

Idp3p Is Required for the Peroxisomal Degradation of Unsaturated Fatty Acids-In search for the function of Idp3p in peroxisomal metabolism, we tested the growth abilities of  $idp3\Delta$  cells on different carbon sources. Cells grew normally on medium containing glucose, glycerol, or stearate as a single carbon source (Fig. 11A). Also, on oleic acid plates, no significant growth differences between wild-type and  $idp3\Delta$  mutant cells were observed (Fig. 11B). In liquid oleic acid medium, however, the generation time of  $idp3\Delta$  mutant cells increased from 8 h as determined for the wild type to 12 h for the mutant (Fig. 11B). Because the only difference between stearic acid and oleic acid is the presence of one double bond in position 9, the observed growth defect suggested that the peroxisomal Idp3p might play a role in the degradation of unsaturated fatty acids. This assumption was further supported by the complete inability of cells lacking Idp3p to grow on petroselinic acid, an unsaturated fatty acid that contains a double bond at position 6 (Fig. 11C). The observed growth defects on oleic acid and petroselinic acid medium were complemented upon transformation of the  $idp3\Delta$  mutant with the wild-type IDP3 gene (Fig. 11, B and C). These results confirmed that the impaired growth of  $idp3\Delta$  mutant cells on unsaturated fatty acids was indeed caused by the lack of Idp3p.

Yeast Peroxisomes Contain Auxiliary Enzymes Needed for the Degradation of Unsaturated Fatty Acids—The ability of S. cerevisiae to grow on unsaturated fatty acids as the single carbon source (Fig. 11), the presence of an NADP-dependent isocitrate dehydrogenase in peroxisomes (Fig. 6), and its suggested role of supplying NADPH for the degradation of unsaturated fatty acids encouraged us to search for auxiliary enzymes of this pathway. Both the  $\Delta 2,\Delta 3$ -encyl-CoA isomerase and the NADPdependent 2,4-dienoyl-CoA reductase activities were detected in whole-cell yeast lysates (data not shown). For subcellular localization of the activities, wild-type yeast homogenates were subjected to sucrose density gradient centrifugation, which did result in a clear separation of peroxisomes and mitochondria as judged by organelle-specific marker enzymes (Fig. 12). Both the  $\Delta 2, \Delta 3$ -enoyl-CoA isomerase and the NADP-dependent 2,4-dienoyl-CoA reductase activities co-segregated with the peroxisomal marker catalase, demonstrating that both enzymes are localized in peroxisomes of S. cerevisiae (Fig. 12). These data

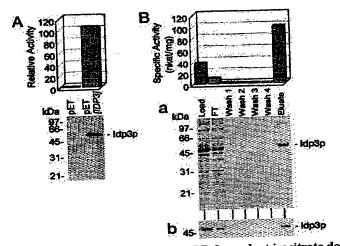


FIG. 8. Idp3p is a peroxisomal NADP-dependent isocitrate de hydrogenase. A, cell homogenates from E. coli BL21(DE3) transformed with the pET vector or pET-IDP3 were analyzed for the presence of Idp3p and of NADP-dependent isocitrate dehydrogenase activity. Equal amounts of cell homogenates were subjected to immunoblot analysis with Idp3p antibodies and to enzyme measurements. Immunoreactive Idp3p was only detected in E. coli transformants expressing IDP3, and expression correlated with the appearance of specific NADP-dependent isocitrate dehydrogenase activity in the bacterial extracts. B, HISe-tagged Idp3p from bacterial extracts of E. coli BL21(DE3) was purified by affinity chromatography on nickel-nitrilotriacetic acid resin (see "Experimental Procedures"). The block diagram shows the specific activity of NADP-dependent isocitrate dehydrogenase of equal portions of the bacterial homogenate (load), the column flow-through (FT), four steps of washing (washes 1-4), and the eluate. Equal portions of the fractions were processed for SDS-PAGE and Coomassie staining (a) and immunoblot analysis with antibodies against the Idp3p (b). As judged by SDS-PAGE and immunoblot analysis, purification of Idp3p was to apparent homogeneity, and the purified protein retained the NADP-dependent isocitrate dehydrogenase activity. nkat, nanokatal.

Fig. 9. Kinetic properties of Idp3p. Idp3p activities at different NADP (A) and isocitrate (B) concentrations are shown as Lineweaver-Burk plots. The resulting apparent  $K_m$  values for NADP and isocitrate are 0.02 mm and 0.05 mm, respectively.

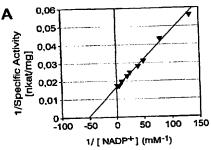
S. cerevisiae harbor the entire suggest that peroxison enzyme equipment needed for the utilization of unsaturated fatty acids, including an NADP-dependent isocitrate dehydrogenase, a putative component of an NADPH-regenerating system.

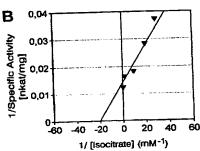
#### DISCUSSION

Here we report on the molecular identification and functional characterization of a peroxisomal NADP-dependent isocitrate dehydrogenase of S. cerevisiae. In line with a role in the peroxisomal metabolism of unsaturated fatty acids, the Idp3p has been demonstrated to be exclusively peroxisomal, and the protein was shown to be essential for the growth of S. cerevisiae on unsaturated fatty acids but dispensable for growth on saturated fatty acids. The supposed function of the protein in peroxisomal fatty acid metabolism is the regeneration of NADPH that is needed by the NADPH-dependent 2,4dienoyl-CoA reductase for the reductive elimination of double bonds of unsaturated fatty acids. This reductase and the  $\Delta 2,\! \Delta 3\! -\!$ encyl-CoA isomerase, another auxiliary enzyme needed for the degradation of unsaturated fatty acids, have been localized to yeast peroxisomes (Fig. 12). The presence of these enzyme activities in peroxisomes has far reaching implications for our understanding of the peroxisomal metabolism and transport of metabolites across the peroxisomal membrane. The data presented are consistent with the assumption that peroxisomes of S. cerevisiae maintain the entire enzyme equipment needed for the degradation of unsaturated fatty acids, including an NADPdependent isocitrate dehydrogenase, a putative constituent of a peroxisomal NADPH-regenerating redox system.

Idp3p was isolated from peroxisomes of oleic acid-induced yeast cells (Fig. 2), and peptide sequence data of the protein were instrumental in cloning the corresponding gene from a genomic yeast library (Fig. 3). Idp3p is exclusively localized in peroxisomes, and consistent with its function in peroxisomal fatty acid metabolism, Idp3p was highly induced upon growth on oleic acid (Figs. 5-7). Protease protection data suggested that the protein resides in the peroxisomal lumen (Fig. 10), which is further supported by the observation that peroxisomal targeting of Idp3p depends on the presence of a C-terminal type 1 peroxisomal targeting signal (data not shown), known to target proteins from the cytosol across the peroxisomal membrane barrier into the peroxisomal matrix (41, 45, 46). That the peroxisomal Idp3p indeed is an NADP-dependent isocitrate dehydrogenase was confirmed by the characterization of the enzymatic properties of purified, recombinant Idp3p (Fig. 8). Interestingly, an NADP-dependent isocitrate dehydrogenase has also been detected in peroxisomes of the n-alkane-utilizing yeast C. tropicalis (44).

Beside the peroxisomal Idp3p, three yeast isoenzymes of isocitrate dehydrogenase have been described that catalyze the oxidative decarboxylation of isocitrate to  $\alpha$ -ketoglutarate. The NAD-specific mitochondrial isoenzyme is an octamer of two nonidentical subunits designated Idh1p and Idh2p (47, 48) and is believed to catalyze a key regulation step in the tricarbonic





acid cycle. Less clear are the function of the two NADP-specific isoenzymes located in mitochondria and the cytosol (17, 18). The glutamate auxotrophy upon deletion of both Idp1p and Idh1p suggest that both enzymes contribute to the anaplerotic supply of  $\alpha$ -ketoglutarate for glutamate formation (40). Furthermore, as isocitrate and  $\alpha$ -ketoglutarate can traverse the mitochondrial membrane via specific transporters (49), it has been suggested that the proteins may participate in an inter-

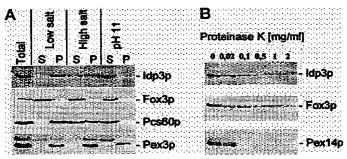


Fig. 10. Subperoxisomal localization of Idp3p. A, extraction of peroxisomes. 25,000 × g organelle pellets were prepared from oleic acid-induced wild-type cells and extracted by low salt, high salt, and carbonate treatment at pH 11.0. Extracted proteins were separated from the membranes by centrifugation. Equal proportions of pellet (P) and supernatant (S) fractions were analyzed by SDS-PAGE and West ern blot analysis using specific antibodies against Idp3p, the integral membrane protein Pex3p (34), Pcs60p (32), and the peroxisomal matrix marker Fox3p (30). As Idp3p is hardly extracted by high salt treatment and totally extracted at pH 11, it behaves like Pcs60p, which is localized in the peroxisomal matrix but is also found loosely associated with the peroxisomal membrane (32). B, protease protection analysis of purified peroxisomes. A cell homogenate of wild-type cells was separated by sucrose density gradient centrifugation, and peroxisomal peak fractions were pooled. Equal amounts of the pooled peroxisomal fractions were incubated in the presence or absence of proteinase K for 10 min on ice. Samples were analyzed by SDS-PAGE and Western blot analysis using antibodies against Idp3p, Pex14p, and Fox3p. The resistance of Idp3p against externally added protease suggests that the protein is protected by the peroxisomal membrane and thus resides in the peroxisomal lumen. In contrast, Pex14p, a peripheral membrane protein localized at the cytosolic face of the peroxisomal membrane (33), is rapidly degraded under these conditions.

ange of reducing equivalents (18). This compartmental raises the question of whether Idp3p might play a comparable role in the peroxisomal metabolism. In the cytosol, NADPH is generated by, for instance, the pentose phosphate pathway. However, because of the impermeability of the peroxisomal membrane for pyrimidine nucleotides (15), the cytosolic NADPH pool cannot directly account for the peroxisomal need for NADPH. This emphasizes the necessity for an NADPHregenerating system in the peroxisomal lumen. Because the formation of  $\alpha$ -ketoglutarate for the production of glutamate is primarily catalyzed by the yeast mitochondrial NAD-dependent and NADP-dependent isocitrate dehydrogenases (40), the most likely biological function of Idp3p is the regeneration of NADPH. The involvement of Idp3p in the intraperoxisomal regeneration of NADPH, which is necessary for the degradation of unsaturated fatty acids, is also more in agreement with the peroxisomal localization and with the oleic acid inducibility of the protein.

The requirement of the peroxisomal degradation of fatty acids with even-numbered double bonds for NADPH is well established (8, 12). Degradation of these fatty acids in the  $\beta$ -oxidation spiral leads to 2,4-dienoyl-CoA intermediates that are reduced to 3-trans-enoyl-CoA in a redox reaction that requires NADPH and that is catalyzed by the peroxisomal NADPH-dependent 2,4-dienoyl-CoA reductase. The resulting 3-trans-enoyl-CoA is subsequently isomerized to 2-trans-enoyl-CoA, which can be reintroduced into the  $\beta$ -oxidation spiral (Fig. 1). The assumption that Idp3p provides the NADPH for this chain of reactions is supported by the observation that cells lacking the protein grow normally on stearate (C18:0) but have lost the ability to grow on petroselinic acid ( $\Delta$ 6-C18:1; Fig. 11).

Until recently, it was generally believed that unsaturated fatty acids with double bonds extending from odd-numbered carbon atoms are chain-shortened to 3-cis-enoyl-CoA esters, which after isomerization to 2-trans-enoyl-CoA are further degraded by the  $\beta$ -oxidation spiral (Fig. 1B) (11). According to this pathway, NADPH would not be needed for the metabolization of these unsaturated fatty acids (Fig. 1). However, Tserng and Jin (50) reported that in mammalian cells also the degradation

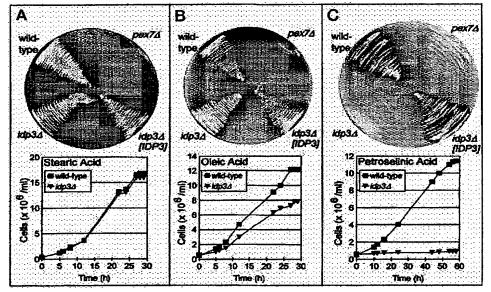
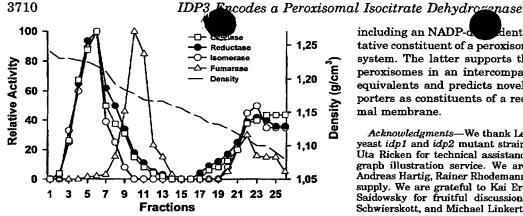


Fig. 11. The peroxisomal NADP-dependent isocitrate dehydrogenase Idp3p is required for growth on unsaturated fatty acids. Growth kinetics from wild-type and  $idp3\Delta$ ,  $idp3\Delta[IDP3]$ , and  $pex7\Delta$  mutant cells on solid agar plates (upper panels) as well as wild-type and  $idp3\Delta$  cells on liquid media (lower panels) containing stearic acid (A), oleic acid (B), or petroselinic acid (C) as a single carbon source. The growth of cells lacking the Idp3p was severely affected on oleic acid media and completely impaired on petroselinic acid-containing media. Growth abilities of  $idp3\Delta$  mutant cells complemented with the IDP3 gene were indistinguishable from those of the wild type.



j,

Fig. 12. Yeast peroxisomes harbor the auxiliary enzymes 2,4dienoyl-CoA reductase and enoyl-CoA isomerase for the degradation of unsaturated fatty acids. Fractions were obtained by isopycnic 20-53% (w/w) sucrose density gradient centrifugation of cell-free extracts from oleic acid-induced wild-type cells. Relative amounts of the peroxisomal marker enzyme catalase and mitochondrial fumarase as well as 2,4-dienoyl-CoA reductase and enoyl-CoA isomerase were monitored. Peroxisomes peaked in fraction 6 at a density of 1.21 g/ml, and mitochondria peaked in fraction 11 at a density of 1.18 g/ml. 2,4-Dienoyl-CoA reductase and enoyl-CoA isomerase activities co-segregated with the peroxisomal marker enzymes, suggesting that both enzymes are localized in peroxisomes.

of unsaturated fatty acids with double bonds extending from odd-numbered carbon atoms requires NADPH. This observation gained support by the exploration of a novel pathway for the reductive removal of odd-numbered double bonds of fatty acids (Fig. 1C) (13). According to this pathway, a  $\Delta 3, 5, \Delta 2, 4$ dienoyl-CoA isomerase, together with the NADPH-dependent 2,4-dienoyl-CoA reductase and the Δ3,Δ2-dienoyl-CoA isomerase facilitate the reduction of odd-numbered double bonds as illustrated in Fig. 1. Recently, it has been suggested that this novel pathway might also be responsible for the degradation of odd-numbered double bonded fatty acids in mammalian peroxisomes (8, 14). In this respect, it is interesting to note that also yeast cells lacking Idp3p are less capable than the wild type of growing on oleic acid (A9-C18:1) as the single carbon

The peroxisomal localization of the Idp3 leads to questions about the origin of the isocitrate and the fate of the a-ketoglutarate that is produced. The most simple explanation would be that α-ketoglutarate is exported directly in exchange for isocitrate as has been demonstrated for mitochondria (49). In principle, isocitrate could also form in peroxisomes from the citrate that is generated by the fusion of acetyl-CoA with oxalacetate, catalyzed by the peroxisomal citrate synthase (Cit2p) (15, 51). However, despite efforts, an aconitase activity has not yet been detected in yeast peroxisomes, thus making the peroxisomal formation of isocitrate from citrate rather unlikely. The direct import of isocitrate from the cytosol into the peroxisomal lumen would predict the existence of a peroxisomal membrane transporter for isocitrate; however, experimental evidence for such a transporter is still missing. In general, our knowledge on the influx and efflux of peroxisomal metabolites and especially on the nature of the carriers involved is still rather limited. For S. cerevisiae, only two peroxisomal metabolite carriers have been described. The heterodimeric Pat1p/ Pat2p ABC-transporter has been suggested to participate in the peroxisomal import of acyl-CoA esters (52), and the peroxisomal carnitin acetyl transferase is involved in the export of the  $\beta$ -oxidation-derived acetyl-CoA (15, 53).

The data presented here are consistent with the idea that peroxisomes of S. cerevisiae maintain the entire enzyme equipment needed for the degradation of unsaturated fatty acids,

including an NADP-d dent isocitrate dehydrogenase, a putative constituent of a peroxisomal NADPH-regenerating redox system. The latter supports the notion of an involvement of peroxisomes in an intercompartmental exchange of reducing equivalents and predicts novel peroxisomal metabolite transporters as constituents of a redox shuttle across the peroxisomal membrane.

Acknowledgments—We thank Lee McAlister-Henn for providing the yeast idp1 and idp2 mutant strains. We thank Ulrike Freimann and Uta Ricken for technical assistance and Siegrid Wüthrich for photograph illustration service. We are grateful to Wolf-Hubert Kunau, Andreas Hartig, Rainer Rhodemann, and Ursula Dorpmund for reagent supply. We are grateful to Kai Erdmann, Peter Rehling, and Jürgen Saidowsky for fruitful discussions. We thank Gabi Dodt, Michael Schwierskott, and Michael Linkert for reading of the manuscript.

#### REFERENCES

- Tolbert, N. E. (1981) Annu. Rev. Biochem. 50, 133-157
   van den Bosch, H., Schutgens, R. B. H., Wanders, R. J. A., and Tager, J. M. (1992) Annu. Rev. Biochem. 61, 157-197
- (1992) Annu. Kev. Biochem. 61, 157-197
   Opperdoes, F. R. (1988) Trends Biochem. Sci. 13, 255-260
   Lazarow, P. B., and Moser, H. W. (1995) in The Metabolic Bases of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) 7th Ed., pp. 2287-2324, McGraw-Hill Inc., New York
   Cooper, T. G., and Beevers, H. J. (1969) J. Biol. Chem. 244, 3514-3520
   Lazarow, P. B., and DeDuve, C. (1976) Proc Natl. Acad. Sci. U. S. A. 73, 2042-2046.
- 2043-2046
- 7. Kunau, W.-H., Bühne, S., de la Garza, M., Kionka, C., Mateblowski, M., de la Garza, M., Schulz-Borchard, U., and Thieringer, R. (1988) Biochem. Soc. Trans. 16, 418-420
- Kunau, W.-H., Dommes, V., and Schulz, H. (1995) Prog. Lipid Res. 34, 267–342
   Hashimoto, T. (1982) Ann. N. Y. Acad. Sci. 386, 5–12
   Veenhuis, M., Mateblowski, M., Kunau, W.-H., and Harder, W. (1987) Yeast 3,

- Stoffel, W., and Caesar, H. (1965) Hoppe-Seyler's Z. Physiol. Chem. 341, 76-83
   Kunau, W.-H., and Dommes, P. (1978) Eur. J. Biochem. 91, 533-544
   Smeland, T. E., Nada, M., Cuebas, D., and Schulz, H. (1992) Proc. Natl. Acad.
- Sci. U. S. A. 89, 6673-6677
   He, X.-Y., Shoukry, K., Chu, C., Yang, J., Sprecher, H., and Schulz, H. (1995)
   *Biochem. Biophys. Res. Commun.* 215, 15-22
   van Roermund, C. W. T., Elgersma, Y., Singh, N., Wanders, J. A., and Tabak,
   H. F. (1995) *EMBO J.* 14, 3480-3486
   Marzioch, M., Erdmann, R., Veenhuis, M., and Kunau, W.-H. (1994) *EMBO J.*
- 13, 4908-4918
- 17. Haselbeck, R. J., and McAlister-Henn, L. (1991) J. Biol. Chem. 266, 2339-2345
  18. Loftus, T. M., Hall, L. V., Anderson, S. L., and McAlister-Henn, L. (1994)
  Biochemistry 33, 9661-9667
  19. Erdmann, R., Veenhuis, M., Mertens, D., and Kunau, W.-H. (1989) Proc. Natl.
  Acad. Sci. U. S. A. 86, 5419-5423
- 20. Erdmann, R., Wiebel, F. F., Flessau, A., Rytka, J., Beyer, A., Fröhlich, K. U.,
- and Kunau, W.-H. (1991) Cell 64, 499-510

  21. Ausubel, F. J., Brent, R., Kingston, R. E., Moore, D. D., Seidmann, J. G., Smith, J. A., and Struhl, K. (1989) Current Protocols in Molecular Biology, Green Publishing Associates and Wiley-Interscience, New York
- Yaffe, M. P., and Schatz, G. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 4819-4823

- Gietz, R. D., and Sugino, A. (1988) Gene (Amst.) 74, 527-534
   Erdmann, R., and Blobel, G. (1995) J. Cell Biol. 128, 509-523
   Fernandez, J., DeMott, M., Atherton, D., and Mische, S. M. (1992) Anal. Biochem. 201, 569-573
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 9.47-9.57, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Daignan-Fornier, B., Nguyen, C. C., Reisdorf, P., Lemeignan, B., and Bolotin-Fukuhara, M. (1994) Mol. Gen. Genet. 243, 575-583
- 28. Wach, A., Brachat, A., Poehlmann, R., and Philippsen, P. (1994) Yeast 10, 1793-1808
- 29. Harlow, E., and Lane, D. (1988) Antibodies, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 30. Erdmann, R., and Kunau, W.-H. (1994) Yeast 10, 1173–1182 31. Rose, M. D., Misra, L. M., and Vogel, J. P. (1989) Cell 57, 1211–1221

- Blobel, F., and Erdmann, R. (1996) Eur. J. Biochem. 240, 468-476
   Albertini, M., Rehling, P., Erdmann, R., Girzalsky, W., Kiel, J. A. K. W., Veenhuis, M., and Kunau, W.-H. (1997) Cell 89, 83-92
- 34. Höhfeld, J., 1167-1178 Veenhuis, M., and Kunau, W.-H. (1991) J. Cell Biol. 114,
- 35. Sikorski, R. S., and Hieter, P. (1989) Genetics 122, 19-27
- 36. Erdmann, R. (1994) Yeast 10, 935-944 37. Erdmann, R., and Blobel, G. (1996) J. Cell. Biol. 135, 111-121
- 38. Moreno de la Garca, M., Schultz-Borchard, U., Crabb, J. W., and Kunau, W.-H. (1985) Eur. J. Biochem. 148, 285-291
- Binstock, J. F., and Schulz, H. (1981) Methods Enzymol. 71, 403-411
   Haselbeck, R. J., and McAlister-Henn, L. (1993) J. Biol. Chem. 268,
- Gould, S. J., Keller, G.-A., and Subramani, S. (1987) J. Cell Biol. 105, 2923-2931
- 42. Elgersma, Y., Vos, A., van den Berg, M., van Roermund, C. W., van der Sluijs,

ij

- P., Distel, B., and Tabak, H. F. (19 diol. Chem. 271, 26375-26382 43. Rottensteiner, H., Kal, A. J., Filipits, M., Binder, M., Hamilton, B., Tabak, H. F., and Ruis, H. (1996) EMBO J. 15, 2924-2934 44. Yamamoto, S., Atomi, H., Ueda, M., and Tanaka, A. (1995) Arch. Microbiol. 163, 104-111 45. Subremeni S. (1993) Annu. Rev. Cell. Rick 9, 445-478
- 163, 104-111
  45. Subramani, S. (1993) Annu. Rev. Cell. Biol. 9, 445-478
  46. McNew, J. A., and Goodman, J. M. (1996) Trends Biochem. Sci. 21, 54-58
  47. Keys, D. A., and McAlister-Henn, L. (1990) Bacteriology 172, 4280-4287
  48. Cupp, J. R., and McAlister-Henn, L. (1991) J. Biol. Chem. 266, 22199-22205

- LaNoue, K. F., and S. J. (1991) J. Biol. Chem. 266, 11614-11620
   Tserng, K. Y., and S., S. J. (1991) J. Biol. Chem. 266, 11614-11620
   Lewin, A. S., Hines, V., and Small, G. M. (1990) Mol. Cell. Biol. 10, 1399-1405
   Hettema, E. H., van Roermund, C. W., Distel, B., van den Berg, M., Vilela, C., Rodrigues-Pousada, C., Wanders, R. J., and Tabak, H. F. (1996) EMBO J. 15, 2813-2829
- 15, 3813-3822 53. Elgersma, Y., van Roermund, C. W., Wanders, R. J., and Tabak, H. F. (1995) *EMBO J.* 14, 3480-3486 54. Struhl, K. (1987) *Cell* 49, 295-297

BNSDOCID: <XP\_\_\_2157778A\_\_I\_>



ELTORGANISATION FUR GEISTIGES EIGENTUM Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6:

C12P 25/00, C12N 15/60, 15/31

(11) Internationale Veröffentlichungsnummer:

WO 97/03208

A1

(43) Internationales Veröffentlichungsdatum:

30. Januar 1997 (30.01.97)

(21) Internationales Aktenzeichen:

PCT/EP96/03009

(22) Internationales Anmeldedatum:

10. Juli 1996 (10.07.96)

(81) Bestimmungsstaaten: CA, CN, JP, KR, US, europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE,

IT, LU, MC, NL, PT, SE).

(30) Prioritätsdaten:

195 25 281.0 195 45 468.5

DE 13. Juli 1995 (13.07.95) DE

6. December 1995 (06.12.95)

Veröffentlicht

Mit internationalem Recherchenbericht.

Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist. Veröffentlichung wird wiederholt falls Änderungen eintreffen.

(71) Anmelder (für alle Bestimmungsstaaten ausser US): BASF AKTIENGESELLSCHAFT [DE/DE]; D-67056 Ludwigshafen (DE). FORSCHUNGSZENTRUM JÜLICH GMBH [DE/DE]; D-52428 Jülich (DE).

(72) Erfinder; und

- (75) Erfinder/Anmelder (nur für US): KÄSLER, Bruno [DE/DE]; Magdeburger Strasse 72, D-67071 Ludwigshafen (DE). SAHM, Hermann [DE/DE]; Wendelinusstrasse 71, D-52428 Jülich (DE). STAHMANN, Klaus-Peter [DE/DE]; Wilhelmstrasse 11, D-52428 Julich (DE). SCHMIDT, Georg [DE/DE]; Heerstrasse 10, D-52457 Aldenhoven (DE). BÖDDECKER, Theo [DE/DE]; Robert-Koch-Strasse 7, D-52428 Jülich (DE). SEULBERGER, Harald [DE/DE]; Adalbert-Stifter-Strasse 4, D-69221 Dossenheim (DE).
- (74) Gemeinsamer Vertreter: BASF AKTIENGESELLSCHAFT; D-67056 Ludwigshafen (DE).
- (54) Title: RIBOFLAVIN-PRODUCTION PROCESS BY MEANS OF MICRO-ORGANISMS WITH MODIFIED ISOCITRATLY ASE **ACTIVITY**
- (54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG VON RIBOFLAVIN MITTELS MIKROORGANISMEN MIT VERÄNDERTER ISOCITRATLYASE-AKTIVITÄT
- (57) Abstract

A microbial riboflavin-production process is disclosed. Riboflavin-producing micro-organisms are cultivated in a culture medium and the thus produced riboflavin is then isolated. The process is characterised in that the endogenous isocitratlyase activity (ICL) of the micro-organisms used has been modified.

## (57) Zusammenfassung

Verfahren zur mikrobiellen Herstellung von Riboflavin durch Kultivierung von Riboflavin produzierenden Mikroorganismen in einem Nährmedium und anschließender Isolierung des hergestellten Riboflavins, dadurch gekennzeichnet, daß Mikroorganismen verwendet werden, bei denen die endogene Isocitratlyase (ICL) Aktivität verändert wurde.

BNSDOCID: <WO\_\_\_9703208A1\_I\_>

# LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AM	Armenien	GB	Vereinigtes Königreich	MX	Mexiko
AT	Österreich	GE	Georgien	NE	Niger
ΑU	Australien	GN	Guinea	NL	Niederlande
BB	Barbados	GR	Griechenland	NO	
BE	Belgien	HU	Ungarn	NZ	Norwegen Neuseeland
BF	Burkina Faso	IE	Irland	PL	Polen
BG	Bulgarien	IT	Italien	PT	
BJ	Benin	JP	Japan	RO	Portugal Rumānien
BR	Brasilien	KE	Kenya	RU	
BY	Belarus	KG	Kirgisistan	SD	Russische Föderation Sudan
CA	Kanada	KP	Demokratische Volksrepublik Korea	SE SE	Sudan Schweden
CF	Zentrale Afrikanische Republik	KR	Republik Korea	SG	
CG	Kongo	KZ	Kasachstan	SI	Singapur
CH	Schweiz	ü	Liechtenstein	SK	Slowenien
CI	Côte d'Ivoire	LK	Sri Lanka		Slowakei
CM	Kamerun	LR	Liberia	SN SZ	Senegal
CN	China	LK	Litauen	SZ TD	Swasiland
CS	Tschechoslowakei	LU	Luxemburg		Tschad
CZ	Tschechische Republik	LV	Lettland	TG	Togo
DE	Deutschland	MC	Monaco	TJ	Tadschikistan
DK	Dänemark	MD	Republik Moldau	TT	Trinidad und Tobago
EE	Estland	MG	Madagaskar	UA	Ukraine
ES	Spanien	ML	Mali	UG	Uganda
FI	Finnland	MN	Mongolei	US	Vereinigte Staaten von Amerika
FR	Frankreich	MR	Mauretanien	UZ	Usbekistan
GA	Gabon	MW	Malawi	VN	Vietnam
		141 44	Mialawi		

Verfahren zur Herstellung von Riboflavin mittels Mikroorganismen mit veränderter Isocitratlyase-Aktivität

## 5 Beschreibung

Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von Riboflavin mittels Mikroorganismen mit veränderter Isocitratlyase-Aktivität.

10

Das Vitamin  $B_2$ , auch Riboflavin genannt, ist für Mensch und Tier essentiell. Bei Vitamin- $B_2$ -Mangel treten Entzündungen der Mundund Rachenschleimhäute, Risse in den Mundwinkeln, Juckreiz und Entzündungen in den Hautfalten u.ä. Hautschäden, Bindehautentzün-

- 15 dungen, verminderte Sehschärfe und Trübung der Hornhaut auf. Bei Säuglingen und Kindern können Wachstumsstillstand und Gewichtsabnahme eintreten. Das Vitamin  $B_2$  hat daher wirtschaftliche Bedeutung insbesondere als Vitaminpräparat bei Vitaminmangel sowie als Futtermittelzusatz. Daneben wird es auch als Lebensmittelfarb-
- 20 stoff, beispielsweise in Mayonnaise, Eiscreme, Pudding etc., eingesetzt.

Die Herstellung von Riboflavin erfolgt entweder chemisch oder mikrobiell. Bei den chemischen Herstellungsverfahren wird das 25 Riboflavin in der Regel in mehrstufigen Prozessen als reines Endprodukt gewonnen, wobei relativ kostspielige Ausgangsprodukte - wie beispielsweise D-Ribose - eingesetzt werden müssen.

Eine Alternative zur chemischen Herstellung des Riboflavins bie30 tet die Herstellung dieses Stoffes durch Mikroorganismen. Als
Ausgangsprodukte für die mikrobielle Synthese können nachwachsende Rohstoffe, wie beispielsweise pflanzliche Öle, eingesetzt
werden.

- 35 Die Herstellung von Riboflavin durch Fermentation von Pilzen wie Ashbya gossypii oder Eremothecium ashbyii ist bekannt (The Merck Index, Windholz et al., eds. Merck & Co., Seite 1183, 1983); aber auch Hefen, wie z.B. Candida oder Saccharomyces, und Bakterien, wie Clostridium, sind zur Riboflavinproduktion geeignet. Ribofla-
- 40 vin-überproduzierende Bakterienstämme sind beispielsweise in der EP 405370 beschrieben, wobei die Stämme durch Transformation der Riboflavin-Biosynthese-Gene aus Bacillus subtilis erhalten wurden. Diese Prokaryonten-Gene sind aber für ein rekombinantes Riboflavin-Herstellungsverfahren mit Eukaryonten wie Saccharomy-
- 45 ces cerevisiae oder Ashbya gossypii ungeeignet.

In WO 93/03183 ist die Klonierung der für die Riboflavin-Biosynthese spezifischen Gene aus dem eukaryontischen Organismus
Saccharomyces cerevisiae beschrieben. Mittels dieser Gene können
rekombinante eukaryontische Mikroorganismen konstruiert werden,
5 die eine effiziente Riboflavinproduktion gestatten.

Häufig liegen jedoch die Ausgangsprodukte und Substrate der Riboflavinbiosynthese-Enzyme in dem Mikroorganimus in limitierter Menge vor, so daß trotz Erhöhung der Riboflavinbiosynthese -10 Aktivität keine Steigerung in der Riboflavinproduktion erreicht wird.

Es bestand daher die Aufgabe ein verbessertes mikrobielles Verfahren zur Produktion von Riboflavin bereitzustellen, das Mikro-15 organismen verwendet, die keine oder eine geringere Substratlimitierung besitzen und somit eine erhöhte Riboflavinproduktion erlauben.

Die Aufgabe wird erfindungsgemäß dadurch gelöst, daß die verwen20 deten Mikroorganismen eine Veränderung in ihrer endogenen Isocitratlyaseaktivität besitzen. Die Veränderung ist gegenüber dem unveränderten Ausgangsstamm zu ermitteln. Es gibt eine Vielzahl von
Möglichkeiten, solche Mikroorganismen mit veränderter ICLAktivität zu erhalten.

25

Eine Möglichkeit besteht darin, das endogene ICL-Gen so zu verändern, daß es für ein Enzym mit gegenüber dem Ausgangsenzym erhöhter ICL-Aktivität codiert. Eine Erhöhung der Enzymaktivität kann beispielsweise erreicht werden, indem durch Veränderung des katalytischen Zentrums ein erhöhter Substratumsatz erfolgt oder indem die Wirkung von Enzyminhibitoren aufgehoben wird. Auch kann eine erhöhte Enzymaktivität durch Erhöhung der Enzymsynthese, beispielsweise durch Genamplifikation oder durch Ausschaltung von Faktoren, die die Enzymbiosynthese reprimieren, hervorgerufen werden. Die endogene ICL-Aktivität wird vorzugsweise durch Mutation des endogenen ICL-Gens erhöht. Derartige Mutationen können entweder nach klassischen Methoden ungerichtet erzeugt werden, wie beispielsweise durch UV-Bestrahlung oder mutationsauslösenden Chemikalien oder gezielt mittels gentechnischer Methoden wie

Die ICL-Genexpression wird durch Erhöhen der ICL-Genkopienzahl und / oder durch Verstärkung regulatorischer Faktoren , die die ICL-Genexpression positiv beeinflussen, erhöht. So kann eine Verstärkung regulatorischer Elemente vorzugsweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und Enhancer verwendet werden. Daneben ist aber auch

eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der m-RNA verbessert wird. Zur Erhöhung der Genkopienzahl wird das ICL-Gen in ein Genkonstrukt bzw. in einen Vektor eingebaut, der vorzugsweise das dem ICL-Gen zugeordnete regulatorische Gensequenzen enthält, insbesondere solche, die die Genexpression verstärken. Anschließend wird ein Riboflavin-produzierenden Mikroorganismus mit dem das ICL-Gen enthaltende Genkonstrukt transformiert.

10 Das ICL-Gen wird vorzugsweise aus Mikroorganismen, insbesondere aus dem Pilz Ashbya gossypii, isoliert. Für die Isolierung des Gens kommen aber auch alle weiteren Organismen, deren Zellen die anaplerotische Sequenz des Glyoxylat-Cyclus und damit die Isocitratlyase enthalten, also auch Pflanzen, in Betracht. Die Isolie-15 rung des Gens kann durch homologe oder heterologe Komplementation einer im ICL-Gen defekten Mutante oder auch durch heterologes Probing oder PCR mit heterologen Primern erfolgen. Zur Subklonierung kann das Insert des komplementierenden Plasmids anschließend durch geeignete Schnitte mit Restriktionsenzymen in 20 der Größe minimiert werden. Nach Sequenzierung und Identifizierung des putativen Gens erfolgt eine paßgenaue Subklonierung durch Fusions-PCR. Plasmide, die die so erhaltenen Fragmente als Insert tragen, werden in die ICL-Gen-defekte Mutante eingebracht, die auf Funktionalität des ICL-Gens getestet wird. Funktionelle 25 Konstrukte werden schließlich zur Transformation eines Riboflavin-Produzenten eingesetzt.

Nach Isolierung und Sequenzierung sind Isocitratlyasegene mit Nukleotidsequenzen erhältlich, die für die in SEQ ID NO:2 angege30 bene Aminosäuresequenz oder deren Allelvariationen kodieren. Allelvariationen umfassen insbesondere funktionelle Derivate, die durch Deletion, Insertion oder Substitution von Nukleotiden aus der in SEQ ID NO:1 dargestellten Sequenz erhältlich sind, wobei die ICL-Aktivität aber erhalten bleibt.

Den Isocitratlyasegenen ist insbesondere ein Promotor der Nukleotidsequenz von Nukleotid 176 bis 550 gemäß SEQ ID NO:1 oder eine im wesentlichen gleichwirkende DNA-Sequenz vorgeschaltet. So kann beispielsweise dem Gen ein Promotor vorgeschaltet sein, der sich von dem Promotor mit der angegebenen Nukleotidsequenz durch ein oder mehrere Nukleotidaustausche, durch Insertion(en) und/oder Deletion(en) unterscheidet, ohne daß aber die Funktionalität bzw. Wirksamkeit des Promotors beeinträchtigt ist. Des weiteren kann der Promotor auch durch Veränderung seiner Sequenz in seiner 45 Wirksamkeit erhöht oder komplett durch wirksamere Promotoren ausgetauscht werden.

Dem ICL-Gen können des weiteren regulatorische Gensequenzen bzw. Regulatorgene zugeordnet sein, die insbesondere die ICL-Gen-Aktivität erhöhen. So können dem ICL-Gen beispielsweise sog. "enhancer" zugeordnet sein, die über eine verbesserte Wechselwirskung zwischen RNA-Polymerase und DNA eine erhöhte ICL-Genexpression bewirken.

Dem Isocitratlyasegen mit oder ohne vorgeschaltetem Promotor bzw. mit oder ohne Regulatorgen können ein oder mehrere DNA-Sequenzen 10 vor- und/oder nachgeschaltet sein, so daß das Gen in einer Genstruktur enthalten ist.

Durch Klonierung des ICL-Gens sind Plasmide bzw. Vektoren erhältlich, die das ICL-Gen enthalten und - wie bereits oben erwähnt 
15 zur Transformation eines Riboflavin-Produzenten geeignet sind.
Die durch Transformation erhältlichen Zellen, bei denen es sich
vorzugsweise um transformierte Zellen von Ashbya gossypii handelt, enthalten das Gen in replizierbarer Form, d.h. in zusätzlichen Kopien auf dem Chromosom, wobei die Genkopien durch homologe

20 Rekombination an beliebigen Stellen des Genoms integriert werden,
und/oder auf einem Plasmid bzw. Vektor.

Eine weitere Möglichkeit, Mikroorganismen mit veränderter ICLAktivität zu erzeugen, besteht darin, Mikroorganismen mit einer

25 Resistenz gegennüber auf ICL hemmend wirkenden Substanzen zu erzeugen und diese zu selektionieren. Hemmstoffe der Isocitratlyase (ICL) sind dem Fachmann bekannt und beispielsweise in Handbook of Enzyme Inhibitors, Herausgeber: Hellmut Zollner, Verlag Chemie, Weinheim, 1993, auf Seite 291 aufgeführt. Besonders geeignete

30 Hemmstoffe sind Phosphoenolpyruvat (PEP), 6-P-Gluconat, Maleat, insbesondere aber Itaconat und Oxalat.

Werden nunmehr Riboflavin produzierende Mikroorganismen-Stämme in Gegenwart solcher Hemmstoffe kultiviert, zeigt sich überraschen35 derweise, daß die Riboflavinbildung gehemmt ist. Dies äußert sich auf Kulturplatten in der Ausbildung von Kolonien, die nicht gelb werden, sondern weiß bleiben. Mit diesem System sind daher Stämme leicht erkennbar, die gegen eine Isocitratlyase-Hemmung resistent sind, da solche Stämme auch in Gegenwart von Hemmstoff Riboflavin bilden und daher gelb gefärbte Kolonien ausbilden. Derartige Stämme können entweder durch Spontanmutation entstehen oder indem entsprechende Mutationen durch gängige Methoden, wie beispielsweise chemisch oder durch UV-Bestrahlung, induziert werden. Es können somit Mikroorganismen-Stämme gewonnen werden, die einen erhöhten Anteil an Riboflavin in das Kulturmedium ausscheiden. Als resistenter Stamm mit erhöhter Riboflavinbildung wurde ins-

besondere der bei der DSM unter der Nr. 10067 hinterlegte Ashbya gossypii-Stamm erhalten.

Als Mikroorganismus werden in dem erfindungsgemäßen Verfahren 5 bevorzugt Pilze eingesetzt. Geeignete Pilze sind beispielsweise solche, die in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) auf Seite 15, Tabelle 6 aufgeführt sind.

Insbesonders sind solche der Gattungen Pichia, Eremothetium und 10 Ashbya, besonders Ashbya gossypii geeignet.

Es können aber auch andere Mikroorganismen als Pilze, beispielsweise Bakterien, insbesondere die, die in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) auf Seite 16, Tabelle 6 aufge-15 führt sind, eingesetzt werden.

Beispiel 1: Erstellung einer genomischen Genbank aus Ashbya gossypii

- 20 Zur Erstellung einer genomischen DNA-Bank wurde chromosomale DNA nach der Methode von Wright und Philippsen (1991, Gene 109: 99-105) isoliert. Die DNA wurde partiell mit Sau 3A verdaut und mit einem Saccharose-Dichtegradienten fraktioniert. Die größten Fragmente (Figur 4) wurden mit dem Bam HI geschnittenen
- 25 E.coli, S.cerevisiae Shuttlevektor YEp 352 (J.E. Hill et al.,1993, Yeast 2: 163-167) ligiert. Mit diesem Ligationsansatz wurde E.coli DH5 a transformiert Von Platten mit Ampicillin und X-Gal wurden 3600 Kolonien isoliert, die durch ihre weiße Farbe als Klone mit Insert tragendem Plasmid erkennbar waren. Die Untersu-
- 30 chung von dreißig solcher zufällig ausgewählter Klone ergab, daß tatsächlich alle ein Plasmid trugen, diese Inserts im Größenbereich 7-18 kb hattten und alle Inserts verschieden waren, was anhand der Restiktionsmuster erkennbar war. Aufgrund einer Genomgröße von 7 x 10<sup>3</sup> kb für Ashbya gossyii liegt die Wahrscheinlich-
- 35 keit, das jedes Gen in dieser Genbank enthalten ist, bei 97 % 99,99 %. Je 100 Klone wurden auf einer Agarplatte in großen Ausstrichen kultiviert und danach die Plasmide als Pool präpariert. Die Genbank bestand dementsprechend aus 36 Plasmidpools.
- 40 Beispiel 2: Selektion des icl1-tragenden Genbankfragments

Mit den Plasmidpräparationen der Genbank wurde die Hefe Saccharomyces cerevisiae ICLld ura3(fs) (E. Fernández et al., 1992, Eur.

**45** J. Biochem. 204: 983-990) transformiert. Diese Mutante ist im ICL1-Gen disruptiert und besitzt im ura3-Gen eine Mutation im Leserahmen. Dieser Genotyp führt dazu, daß der Stamm nicht auf

Ethanol als Kohlenstoffquelle wachsen kann und eine Uracil-Auxotrophie zeigt. Im ersten Schritt wurden die mit der Genbank
transformierten Hefezellen auf Minimalmedium mit Glucose als einziger Kohlenstoffquelle selektioniert. Aufgrund des auf dem

5 Plasmid vorhandenen ura3-Gens konnten nur die Zellen wachsen, die
ein Plasmid aufgenommen hatten, denn das Minimalmedium enthielt
kein Uracil. In diesem Schritt wurden 1900 Klone erhalten. Diese
wurden durch Replikaplattierung auf ein Minimalmedium mit Ethanol
als einziger Kohlenstoffquelle übertragen. Da zum Wachstum auf

10 Ethanol unbedingt die Isocitratlyase als anaplerotisches Enzym nötig ist, konnten nur die Klone wachsen, die auf dem Plasmid das ICL-Gen trugen. Es konnten zwei Klone isoliert werden, die auf Ethanol wuchsen.

## 15 Beispiel 3:

Überprüfung der Funktionalität des isolierten Genbankfragments

Zur Überprüfung, ob die Komplementierung des chromosomalen ICLDefekts plasmid-kodiert war, wurden die selektionierten Saccharo20 myces-Klone zweimal auf Vollmedium mit Uracil kultiviert und die
erhaltenen Zellen auf Platten vereinzelt. Von 16 bzw. 13 zufällig
ausgewählten Klonen wuchsen 7 bzw. 5 nicht mehr auf Minimalmedium
mit Glucose. Genau diese Klone wuchsen auch nicht mehr auf Minimalmedium mit Ethanol. Die Kurierung vom Plasmid war also mit dem
25 Verlust der ICL1d-Komplementation korreliert.

Aus einem der beiden Klone wurde das Plasmid wieder isoliert Es enthielt ein Insert von etwa 8 kb. Erneute Transformation der Saccharomyces. Mutante führte zur Komplementation aller gefunde30 nen Klone. Das 8 kb - Fragment ließ sich durch Sph I auf 2,9 kb, die voll funktionell waren, verkürzen.

Im Rohextrakt der auf Ethanol gewachsenen Transformande war die Isocitratlyase mit einer spezifischen Aktivität von 0,3 U/mg Pro-35 tein meßbar. Zudem zeigte der Westernblott mit polyklonalen Antikörpern gegen die Ashbya-ICL ein deutliches Signal.

PCR mit von tryptischen Peptiden der ICL abgeleiteten Primern ergab starke Signale der erwarteten Größe. Aus einem zweidimensio40 nalen Elektrophoresegel wurde ein Protein isoliert, mit Trypsin

in Peptide zerlegt und durch Edmannabbau ansequenziert. Der Vergleich der Peptidsequenzen mit Datenbanken ergab eine Identität von über 70 % mit der Isocitratlyase aus Saccharomyces cerevisiae. Davon abgeleitete Primer wurden zur PCR eingesetzt.

45 Von dem ca. 8 kb großen komplementierenden Genbankfragement wurden 3,3 kb sequenziert (Sanger et al. Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467). Auf der ermittelten Sequenz konnten durch

7

Datenbankvergleich zwei kodierende Bereiche gefunden werden. Ein Leserahmen von 1680 Basen (SEQ ID NO:1) zeigt eine 65 %ige Identität zum ICL1-Gen von Saccharomyces cerevisiae. Das ICL-Gen liegt 375 Basen upstream von einer Sequenz die 84 % Identität zu 5 einer Ser-tRNA von Saccharomyces cerevisiae zeigt (SEQ ID NO:1).

## Beispiel 4:

Funktionalität subklonierter ICL in einem E.coli/Hefe/Ashbya - Shuttlevektor

10

Zwei durch Restriktionsverdau erhaltene Fragmente und ein PCR-Produkt des isolierten Genbankfragments (Figur 5) wurden in das von Steiner und Philippsen (1994, Mol. Gen. Genet 242: 263-271) konstruierte Plasmid pAG 100 (Figur 6) kloniert. Bei den Fragmen-

- 15 ten handelte es sich um ein 2.9 kb Sph I- Fragment (pAG 100 icl.4) und um ein 2.2 kb Bgl 1 / Eco RV Fragment (pAG 100 icl.6). Beide Fragment enthielten die Ser-tRNA. Deshalb wurde zusätzlich eine PCR-Amplifikation des putativen Gens mit daran fusionierten Bam HI Schnittstellen (pAG 100 icl.8) durchgeführt.
- 20 Alle drei DNAs wurden in die Bam HI site des Plasmids pAG 100 kloniert. Mit den erhaltenen Plasmiden wurde die Hefemutante Saccharomyces cerevisiae ICLld ura3 (fs) transformiert. Alle drei Konstrukte führten zur vollständigen Komplementation der ICLld-Disruption d.h. trugen funktionelle Gene.

25

#### Beispiel 5:

Wirkung der ICL tragenden Plasmide auf die Riboflavinbildung von Ashbya gossypii

- 30 Die Transformation von Ashbya gossypii (Methode: Wright und Philippsen, 1991, Gene 109: 99-105) mit den oben erklärten Plasmiden führte zu signifikanten Erhöhungen der Riboflavinbildung. Kultiviert wurde in 500 ml Schüttelkolben mit zwei Schikanen, das 50 ml Medium aus 10 g/I Sojaöl, 10 g/I Hefeextrakt und 200 µg/ml
- 35 Geneticin enthielt. Der Kontrollstamm A.gossypii pAG 100, der ein Plasmid ohne Insert enthielt, produzierte in zwei Tagen 18,7  $\pm$  0,1 mg/l Riboflavin. Die Stämme A. gossypii pAG 100.4 und Agossypii pAG 100.6 produzierten 31,2  $\pm$  6,1 mg/l bzw. 31,0  $\pm$  2,0 mg/l Riboflavin (Figur 7). Eine signifikante Änderung der spezifischen
- 40 Aktivität der Isocitratlyase war aufgrund der starken Streuung nicht messbar. Der Stamm A. gossypii pAG 100.8 produzierte in einem Medium, das noch durch 3 g/I Glycin supplementiert wurde, innerhalb von drei Tagen 65  $\pm$  5,6 mg/I Riboflavin. Der Kontrollstamm A.gossypii pAG 100 bildete dagegen im direkten Vergleich nur 29,9  $\pm$  1,8 mg/l Riboflavin (Figur 8). Weder in der spezifi-

В

schen Aktivität der Isocitratlyase noch im Myzeltrockengewicht waren signifikante Unterschiede meßbar.

## Beispiel 6:

5 Reinigung einer Isocitratlyase (ICL)

Zur Identifizierung von auf ICL hemmend wirkenden Substanzen wurde zunächst die ICL aus Ashbya gossypii gereinigt. Die Isolierung und Reinigung des Enzyms erfolgte 10nach Wachstum des Pilz10 mycels auf Pflanzenöl. Die einzelnen Reinigungsschritte sind in der Tabelle 1 zusammengefaßt: Demgemäß enthält ein typischer, aus ca. 25 g Mycel hergestellter Rohextrakt, der durch Zellaufschluß mit einer French-Press gewonnen wurde, 220 Einheiten ICL-Aktivität. Etwa 78% davon sind nach Zentrifugation bei 40.000 g gelöst im Überstand wiederzufinden. Eine anschließende fraktionierte Ammoniumsulfatfällung führt zu einer dreifachen Anreicherung des Enzyms. Nach einer Gelfiltration mit einer Sephacryl S-300 Säule wird die TCL an den Kationenaustauscher Mono S-Sepharose gebunden und mit NaCl eluiert. Das so erhaltene Präparat ist homogen in der SDS-Polyacrylamidgelelektrophorese und hat eine spezifische Aktivität von 18,4 U/mg.

#### Beispiel 7:

Identifizierung von ICL-Hemmstoffen

25

Mit dem gereinigten Enzym lassen sich in einem colorimetrischen Test (Dixon, H. und Kornberg, H.L. (1959), Biochem. J. 72, 3: Assay methods for key enzymes of the glyoxylate cycle) Einflüsse von Substanzen auf die Aktivität messen. In Tabelle 2 und Figur 1

- 30 sind die Effekte der getesteten Substanzen auf das Enzym zusammengefaßt bzw. dargestellt. Untersucht wurden zum einen Substanzen, die als Metaboliten in der Pilzzelle einen hemmenden Effekt auf das Enzym haben könnten. Darunter zeigten 6-P-Gluconat und Phosphoenolpyruvat die deutlichsten Hemmwirkungen mit über
- 35 50% bei einer Konzentration von 10 mM. Erheblich besser wirkten jedoch Itakonat und Oxalat, die vermutlich nicht im Stoffwechsel des Pilzes vorkommen. Bereits eine Konzentration von 1 mmol führte zu 78% bzw. 95% Hemmung.

#### 40 Beispiel 8:

Charakterisierung einer mit Itakonat selektionierten Mutante

Durch UV-Bestrahlung von isolierten Sporen des Pilzes lassen sich Mutationen im Erbmaterial erzeugen. Mit einer Strahlendosis, bei 45 der 10-20% der eingesetzten Sporen überleben, erhält man Mutanten, die gegen eine Hemmung der Riboflavinbildung durch Itakonat

resistent sind. Eine so isolierte Mutante zeigt bei Wachstum auf

Sojaöl eine 25-fache Riboflavinbildung im Vergleich zum Ausgangsstamm (Figur 2). Die spezifische ICL-Aktivität ist während der Ribof lavinbildungsphase um bis zu 15% erhöht (Figur 2). Mit Antikörpern läßt sich zeigen, daß die Proteinmenge erhöht ist. Die 5 ICL aus der Mutante zeigt das gleiche Hemmverhalten durch Itakonat wie der Ausgangsstamm.

Beispiel 9:

Korrelation von Riboflavinbildung und spezifischer ICL-Aktivität

Einen überraschenden Hinweis auf einen kausalen Zusammenhang zwi-

schen ICL und Riboflavinbildung liefert die Beobachtung, daß der Pilz, wenn Glucose als Substrat angeboten wird, erst nach Verbrauch der Glucose mit der Produktion beginnt. Genau dann wird

15 auch die ICL, die zuvor durch Glucose reprimiert ist, im Rohextrakt meßbar und steigt bis zu Aktivitäten, wie sie bei Wachstum auf Öl gefunden werden, an (Figur 3).

20

25

30

35

40

45

#### SEQUENZPROTOKOLL

#### (1) ALGEMEINE INFORMATION:

- (i) ANMELDER:
  - (A) NAME: BASF Aktiengesellschaft
  - (B) STRASSE: Carl-Bosch-Strasse 38
  - (C) ORT: Ludwigshafen
  - (E) LAND: Bundesrepublik Deutschland
  - (F) POSTLEITZAHL: D-67056
  - (G) TELEPHON: 0621/6048526
  - (H) TELEFAX: 0621/6043123
  - (I) TELEX: 1762175170
  - (A) NAME: Forschungszentrum Juelich GmbH
  - (B) STRASSE: Leo-Brandt-Strasse
  - (C) ORT: Juelich
  - (E) LAND: Germany
  - (F) POSTLEITZAHL: D-52425
  - (G) TELEPHON: 02461-61 3004
- (ii) ANMELDETITEL: Verfahren zur Herstellung von Riboflavin mittels Mikroorgaismen mit veraenderter Isocitratlyase Aktivitaet
- (iii) ANZAHL DER SEQUENZEN: 2
- (iv) COMPUTER-LESBARE FORM:
  - (A) DATENTRÄGER: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)
- (2) INFORMATION ZU SEQ ID NO: 1:
  - (i) SEQUENZ CHARAKTERISTIKA:
    - (A) LÄNGE: 2364 Basenpaare
    - (B) ART: Nukleins "ure
    - (C) STRANGFORM: Doppel
    - (D) TOPOLOGIE: linear
  - (ii) ART DES MOLEKÄLS: cDNS zu mRNS
  - (iii) HYPOTHETISCH: NEIN
  - (iii) ANTISENSE: NEIN
  - (ix) MERKMALE:
    - (A) NAME/SCHLÜSSEL: 5'UTR
    - (B) LAGE: 1..550
  - (ix) MERKMALE:
    - (A) NAME/SCHLÜSSEL: CDS
    - (B) LAGE: 551..2233

PCT/EP96/03009

11

# (ix) MERKMALE:

WO 97/03208

(A) NAME/SCHLÜSSEL: 3'UTR

(B) LAGE: 2234..2364

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

CGAAAGCGCC AAATACCGGA AACGGCACAG GCGCAGCTCT AATAGCCGTT CCACGATAAC	60
TTTGGAAGTT ATGGCACTAT GGCCGAGTGG TTAAGGCGAC AGACTTGAAA TCTGTTGGGC	120
TCTGCCCGCG CTGGTTCAAA TCCTGCTGGT GTCGTTATTT TTGCCGTTTC TTTTTAGATG	180
AAACTCAGGG GCCTTTAGTC CGCCCTTTTG CCCGCTGATT CATCGCCCGC CAGCAACACC	240
GGTTGAGCCG ATCAGCGCAA GAACGCGCAA AGTCACGTAT GGCCCCTAAG AGTTGAGCTC	300
TCCCCCTCGG CTCCTTCCGG GCGCGGAAAA GCCTGCGTCA CCCCATTAAG TCCGAAACCG	360
CGTTCAAGTG TACTTGGTCC GGGCCAATGT GGTTGCCTCA TCCGAGTCAC CGATACGCAG	420
GTGCGCCCGT CGAGTCACCA TTAGGAGTAG AGCATCTGAT TATATATAGG CCTAGTTACA	480
GCGGTAACAT AGACTGATAG CTCCAGCTCC AGCACTAGCT TGTAGGACAT CTGCGCGACA	540
CCCAGTGAAC ATG TCC CCT TCC GTC AGA GAC GCC CGC AAC GAC CTT GCC Met Ser Pro Ser Val Arg Asp Ala Arg Asn Asp Leu Ala	589
1 5 10	
AGC CTG CAA CAG CAG GCA GCC GCC GAA GCC GAG GAT ATT AGG AGA TGG Ser Leu Gln Gln Ala Ala Ala Glu Ala Glu Asp Ile Arg Arg Trp 15 20 25	637
TGG AGC CAG CCA CGG TGG GCG GGC ACC AAG CGC GTG TAC ACG GCC GAG Trp Ser Gln Pro Arg Trp Ala Gly Thr Lys Arg Val Tyr Thr Ala Glu	685
30 35 40 45	722
GAC ATC GTC AAG CGC CGC GGC ACG TTC CCT GTC GTC GAA TAC CCA TCT Asp Ile Val Lys Arg Arg Gly Thr Phe Pro Val Val Glu Tyr Pro Ser 50 55 60	733
TCC GTA ATG GCG GAC AAG CTC GTG GAG ACA TTG GCG CGG CAC TCG CGC	781
Ser Val Met Ala Asp Lys Leu Val Glu Thr Leu Ala Arg His Ser Arg 65 70 75	
AAC GGC ACG GTT TCA CAG ACG TTC GGA GTG CTC GAC CCA GTG CAA ATG Asn Gly Thr Val Ser Gln Thr Phe Gly Val Leu Asp Pro Val Gln Met 80 85 90	829
ACG CAA ATG GTG AAG TAT CTG GAC ACG ATT TAC GTG TCT GGC TGG CAA	877
Thr Gln Met Val Lys Tyr Leu Asp Thr Ile Tyr Val Ser Gly Trp Gln 95 100 105	
TGC AGC GCC ACG GCT TCG ACC TCG AAC GAG CCT GGG CCC GAT CTC GCG	925
Cys Ser Ala Thr Ala Ser Thr Ser Asn Glu Pro Gly Pro Asp Leu Ala 110 115 120 125	
GAC TAT CCG ATG GAC ACC GTG CCA AAC AAG GTC GAG CAC CTG TTC ATG Asp Tyr Pro Met Asp Thr Val Pro Asn Lys Val Glu His Leu Phe Met	973

12

PCT/EP96/03009

							1.	4								
				130	כ				13	5				14	0	
GCG Ala	CAG Gln	CTC Lev	TTC Phe 145	His	C GAC S Asp	CGG Arg	AA Lys	A CAG S Gli 150	n Arg	C GAO	G GC	C CGG a Arg	C CT g Le 15	u Se	G TGC r Cys	1021
ACT Thr	ACC Thr	CAG Gln 160	Arg	GAC Glu	CTC Leu	GAC Asp	CAZ Glr 165	ı Leı	G GG(	G CCT	r GA	G AT. u Ile 170	e As	С ТА р Ту	C TTG r Leu	1069
Arg	CCG Pro 175	ATT	GTC Val	GCT Ala	GAC Asp	GCA Ala 180	Asp	ACC Thi	C GG(	C CAC	GG( Gl <sub>3</sub> 185	/ Gl	G CT	A AC ı Th	A GCC r Ala	1117
GTC ( Val 190	TTT Phe	AAA Lys	CTC Leu	ACG Thr	AAG Lys 195	ATG Met	TTC Phe	ATC : Ile	GAG Glu	G CGC Arg 200	Gly	r GCA / Ala	A GCC a Ala	C GG	T ATC 7 Ile 205	1165
CAC A	ATG Met	GAG Glu	GAC Asp	CAG Gln 210	TCC Ser	TCC Ser	AGC Ser	AAC Asn	AAA Lys 215	Lys	TG( Cys	GGG Gly	CAC His	E ATO	G GCG Ala	1213
GGC (	CGC Arg	ŤGC Cys	GTG Val 225	ATC Ile	CCT Pro	GTT Val	CAG Gln	GAG Glu 230	His	ATT	AGT Ser	CGT Arg	TTA Leu 235	[Va]	ACT Thr	1261
GTG (	CGC Arg	ATG Met 240	TGT Cys	GCG Ala	GAC Asp	GTG Val	ATG Met 245	CAC His	TCG Ser	AAC Asn	CTG Leu	GTG Val 250	CTT Leu	' GTC Val	GCG Ala	1309
AGA A Arg T	ACA Thr 255	GAC Asp	TCG Ser	GAG Glu	GCC Ala	GCC Ala 260	ACC Thr	TTA Leu	CTT Leu	AGC Ser	TCG Ser 265	AAC Asn	ATT Ile	GAC Asp	GCG Ala	1357
CGC G Arg A 270	TAS Asp	CAT His	TAC Tyr	TAC Tyr	ATT Ile 275	GTC Val	GGG Gly	GCC Ala	TCG Ser	AAC Asn 280	CCT Pro	GAG Glu	GTA Val	ACT Thr	GTA Val 285	1405
CCG C Pro L	TG eu	ATC Ile	GAA Glu	GTT Val 290	TTG Leu	GAC Asp	GCC Ala	GCG Ala	CAG Gln 295	CAG Gln	GCC Ala	GGC Gly	GCC Ala	TCA Ser 300	GGT Gly	1453
GAC A Asp A	.GA	Leu	GCT Ala 305	CAG Gln	CTA Leu	GAG Glu	GAG Glu	GAC Asp 310	TGG Trp	TGC Cys	AAG Lys	AAG Lys	GCC Ala 315	AAG Lys	TTG Leu	1501
AGG C Arg L	eu :	TTC Phe 320	CAC His	GAG Glu	GCA Ala	Phe	GCC Ala 325	GAC Asp	CAG Gln	GTG Val	AAT Asn	GCC Ala 330	AGC Ser	CCT Pro	TCG Ser	1549
ATC A Ile L	AA ( ys 2 35	GAC . Asp :	AAG Lys .	GCG Ala	Gly '	GTT . Val 340	ATT Ile	GCC Ala	AAA Lys	Phe .	AAC Asn 345	TCA Ser	CAG Gln	ATC Ile	GGG Gly	1597
CCA CA Pro GA 350	AG A	ACA ( Phr (	GGC ( Gly /	Ala	TCG A Ser : 355	ATC A	AGA Arg (	GAG Glu	Met .	CGC / Arg 1	AAA Lys	CTG (	GGC Gly	CGC Arg	GAG Glu 365	1645

13

						13									
CTG CTC Leu Leu															1693
GAG GGC Glu Gly	L'eu														1741
GCA CGC Ala Arg															1789
TTC CCT Phe Pro 415															1837
AAG TTC Lys Phe 430															1885
TGG CCG Trp Pro		Ala													1933
CTG GGC Leu Gly	Glu :														1981
CAT ACC										Arg					2029
TTC GGA Phe Gly 495															2077
GAG GGC Glu Gly 510															2125
GTT GAC Val Asp		Ile :										Ser			2173
TCG ATG Ser Met	Gly I					Glu					Ser				2221
GCC AAA Ala Lys		rgat.	ATCA	TC T	CTGA	GTCA	TT T.	CTCT	'CGAC	AAG	ATCC	TCG			2270
GCCAGACT	TC TO	GGAA'	TATA	TA T	'AACA	TCGG	GTA	.cccc	GAC	ATCC	CTGC	CT T	CCGC	AACGT	2330
GCGAAGCA	GC TO	GATA	CGTA	T AC	TTTA	AACG	CAC	Ά							2364

(2) INFORMATION ZU SEQ ID NO: 2:

WO 97/03208 PCT/EP96/03009

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 560 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear
- (ii) 'ART DES MOLEKÜLS: Protein
- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Met Ser Pro Ser Val Arg Asp Ala Arg Asn Asp Leu Ala Ser Leu Gln
1 5 10 15

Gln Gln Ala Ala Glu Ala Glu Asp Ile Arg Arg Trp Trp Ser Gln
20 25 30

Pro Arg Trp Ala Gly Thr Lys Arg Val Tyr Thr Ala Glu Asp Ile Val 35 40 45

Lys Arg Arg Gly Thr Phe Pro Val Val Glu Tyr Pro Ser Ser Val Met 50 55 60

Ala Asp Lys Leu Val Glu Thr Leu Ala Arg His Ser Arg Asn Gly Thr 65 70 75 80

Val Ser Gln Thr Phe Gly Val Leu Asp Pro Val Gln Met Thr Gln Met 85 90 95

Val Lys Tyr Leu Asp Thr Ile Tyr Val Ser Gly Trp Gln Cys Ser Ala 100 105 110

Thr Ala Ser Thr Ser Asn Glu Pro Gly Pro Asp Leu Ala Asp Tyr Pro 115 120 125

Met Asp Thr Val Pro Asn Lys Val Glu His Leu Phe Met Ala Gln Leu 130 135 140

Phe His Asp Arg Lys Gln Arg Glu Ala Arg Leu Ser Cys Thr Thr Gln 145 150 155 160

Arg Glu Leu Asp Gln Leu Gly Pro Glu Ile Asp Tyr Leu Arg Pro Ile 165 170 175

Val Ala Asp Ala Asp Thr Gly His Gly Gly Leu Thr Ala Val Phe Lys
180 185 190

Leu Thr Lys Met Phe Ile Glu Arg Gly Ala Ala Gly Ile His Met Glu 195 200 205

Asp Gln Ser Ser Ser Asn Lys Lys Cys Gly His Met Ala Gly Arg Cys 210 215 220

Val Ile Pro Val Gln Glu His Ile Ser Arg Leu Val Thr Val Arg Met 225 230 235 240

Cys Ala Asp Val Met His Ser Asn Leu Val Leu Val Ala Arg Thr Asp 245 250 255

Ser Glu Ala Ala Thr Leu Leu Ser Ser Asn Ile Asp Ala Arg Asp His 260 265 270

15

							12								
Tyr	Tyr	Ile 275	Val	Gly	Ala	Ser	Asn 280	Pro	Glu	Val	Thr	Val 285	Pro	Leu	Ile
Glu	Val 290	Leu	Asp	Ala	Ala	Gln 295	Gln	Ala	Gly	Ala	Ser 300	Gly	Asp	Arg	Leu
Ala 305	Gln	Leu	Glu	Glu	Asp 310	Trp	Cys	Lys	Lys	Ala 315	Lys	Leu	Arg	Leu	Phe 320
His	Glu	Ala	Phe	Ala 325	Asp	Gln	Val	Asn	Ala 330	Ser	Pro	Ser	Ile	Lys 335	Asp
Lys	Ala	Gly	Val 340	Ile	Ala	Lys	Phe	Asn 345	Ser	Gln	Ile	Gly	Pro 350	Gln	Thr
Gly	Ala	Ser 355	Ile	Arg	Glu	Met	Arg 360	Lys	Leu	Gly	Arg	Glu 365	Leu	Leu	Gly
Gln	Asp 370	Val	Tyr	Phe	Asp	Trp 375	Asp	Leu	Pro	Arg	Ala 380	Arg	Glu	Gly	Leu
Tyr 385	Arg	Tyr	Lys	Gly	Gly 390	Thr	Gln	Cys	Ala	Ile 395	Met	Arg	Ala	Arg	Ala 400
Phe	Ala	Pro	Tyr	Ala 405	Asp	Leu	Val	Trp	Phe 410	Glu	Ser	Asn	Phe	Pro 415	Asp
Phe	Gln	Gln	Ala 420	Lys	Glu	Phe	Ala	Gln 425	Gly	Val	Arg	Glu	Lys 430	Phe	Pro
Asn	Lys	Trp 435		Ala	Tyr	Asn	Leu 440	Ser	Pro	Ser	Phe	Asn 445	Trp	Pro	Lys
Ala	Met 450	Pro	Pro	Lys	Glu	Gln <b>45</b> 5	Glu	Asn	Tyr	Ile	Gln 460	Arg	Leu	Gly	Glu
Ile 465		Tyr	Val	Trp	Gln 470	Phe	Ile	Thr	Leu	Ala 475	Gly	Leu	His	Thr	Asn 480
Ala	Leu	Ala	Ile	Asp 485	Asn	Phe	Ser	Arg	Glu 490	Phe	Ser	Arg	Phe	Gly 495	Met
Arg	n Ala	Tyr	Ala 500		Gly	Ile	Gln	Gln <b>5</b> 05	Arg	Glu	Met	Asp	Glu 510	Gly	Val
Asr	Val	Leu 515		His	Gln	Lys	Trp 520	Ala	Gly	Ala	Glu	Tyr 525	Val	Asp	Ser
Ile	530		. Leu	ı Ala	Gln	Gly 535	Gly	Val	Ser	Ser	Thr 540	Ala	Ser	Met	Gly
Lys 545		val	Thr	Glu	Glu 550		Phe	Gly	Ser	Ser 555	Asn	Gly	Ala	Lys	Leu 560

#### Patentansprüche

 Verfahren zur mikrobiellen Herstellung von Riboflavin durch Kultivierung von Riboflavin produzierenden Mikroorganismen in einem Nährmedium und anschließender Isolierung des hergestellten Riboflavins, dadurch gekennzeichnet, daß Mikroorganismen verwendet werden, bei denen die endogene Isocitratlyase (ICL) Aktivität verändert wurde.

10

- 2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Mikroorganismen durch Mutation des endogenen ICL-Gens ein Enzym mit höherer ICL-Aktivität aufweisen.
- 15 3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Mikroorganismen durch eine Erhöhung der ICL-Genkopienzahl eine höhere ICL-Genexpression besitzen.
- Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß das
   ICL-Gen mit regulatorischen DNA-Sequenzen funktionell verknüpft wurde, die eine verstärkte Genexpression des ICL-Genserlauben.
- 5. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß Mikro25 organismen mit Resistenz gegenüber auf ICL hemmend wirkenden Substanzen verwendet werden.
- Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß die Mikroorganismen resistent gegenüber den Stoffen Itakonat oder
   Oxalat sind.
  - 7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß als Mikroorganismus ein Pilz verwendet wird.
- 35 8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß der Pilz aus der Gattung Ashbya verwendet wird.
  - 9. ICL-Gen codierend für die in SEQ ID NO: 2 dargestellte Aminosäuresequenz.

40

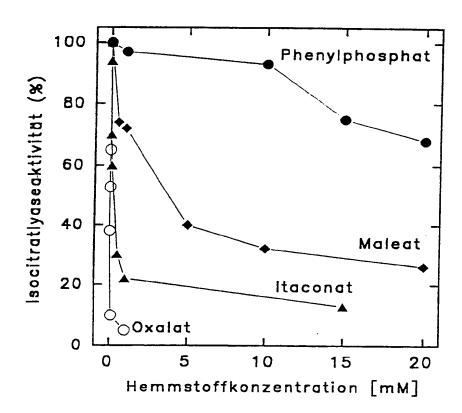
- 10. Genkonstrukt enthaltend ein ICL-Gen gemäß Anspruch 9.
- Genkonstrukt nach Anspruch 10, dadurch gekennzeichnet, daß
  das ICL-Gen funktionell mit einem oder mehreren Regulationssignalen zur Erhöhung der Genexpression funktionell verknüpft
  wurde.

Fraktion	Gesamtaktivität Gesamtprotein	Gesamtprotein	spez. Aktivität	Reinigungsfaktor	Ausbeute
	(Units)	(mg)	(U/mg protein)	(-fach)	(%)
Rohextrakt	220	1310	0.17	1.0	100
40,000 g Überstand	170	730	0.23	1.3	78
35% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> Üherstand	160	630	0.25	1.5	72
60% (NII <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> Pellet	160	300	0.53	3.1	72
Sephacryl S-300 Eluat	52	5	10.8	63	23
Mono S Eluat	35	0.5	18.4	108	16

Tabelle 1.

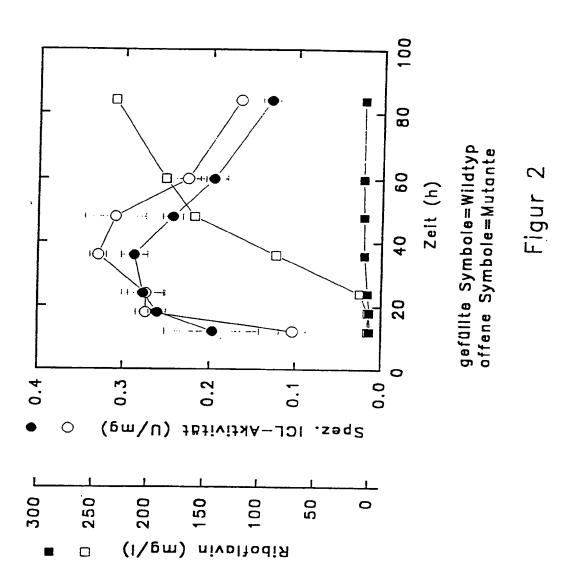
Hemmstoff	Konzentration	Hemmung	Hemmtyp
	(mM)	(%)	
Glucose-6-P	10	\$	
Citrat	10	22	
Fumarat	10	25	
Succinat	10	34	noncompetitive; K <sub>1</sub> :15.8 mM
Malat	10	36	
рбр	10	55	hyperbolic mixed-type
6-P-Gluconat	10	09	
Glycin	10	<5	
Aspartat	10	13	
Glutamat	10	15	
Phenylphosphat	10	7	
Maleat	10	89	
Itaconat		78	linear mixed-type; K <sub>I</sub> :0.17 mM
Oxalat	-	95	noncompetitive; K <sub>i</sub> :0.004 mM

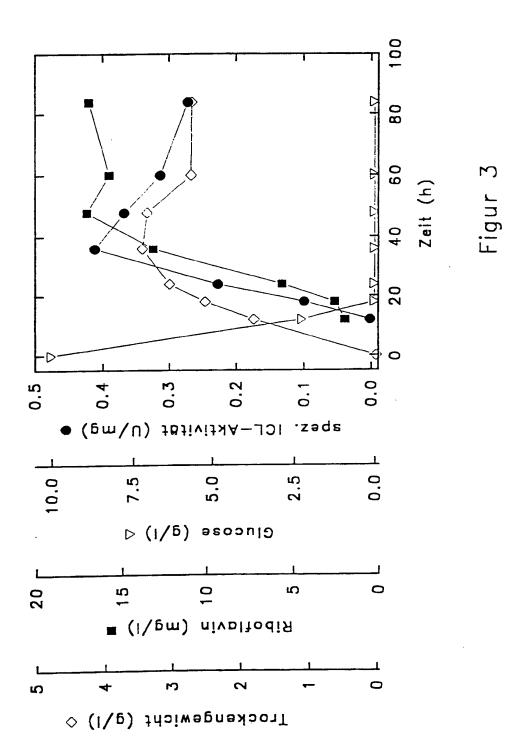
Tabelle 2

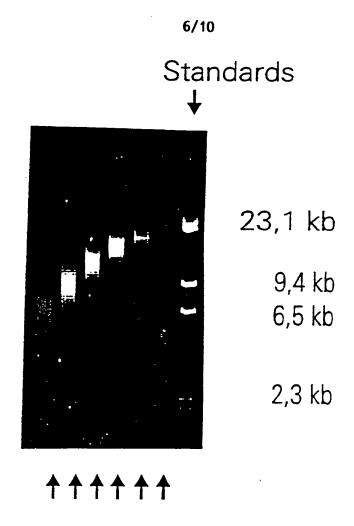


Figur 1

4/10

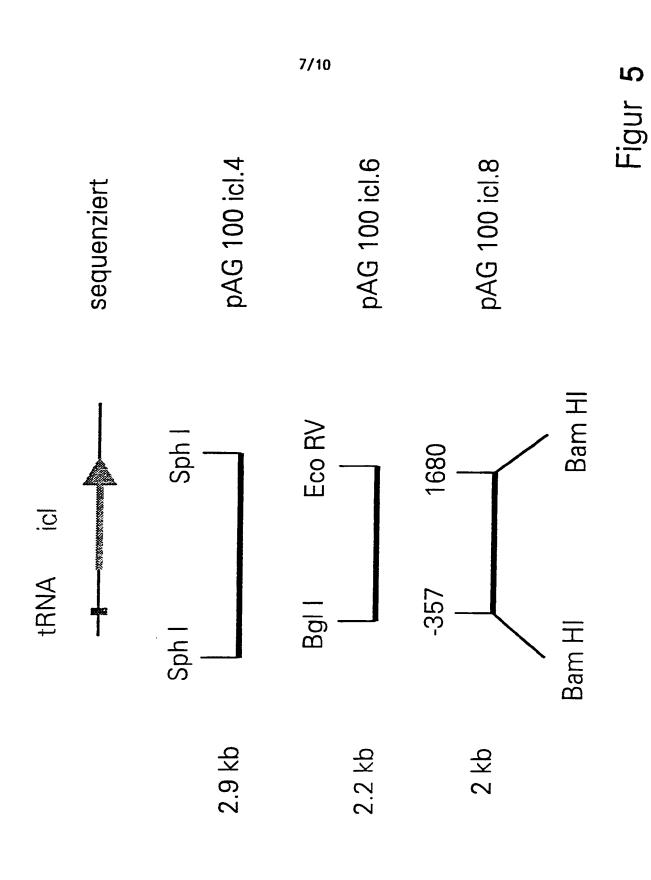






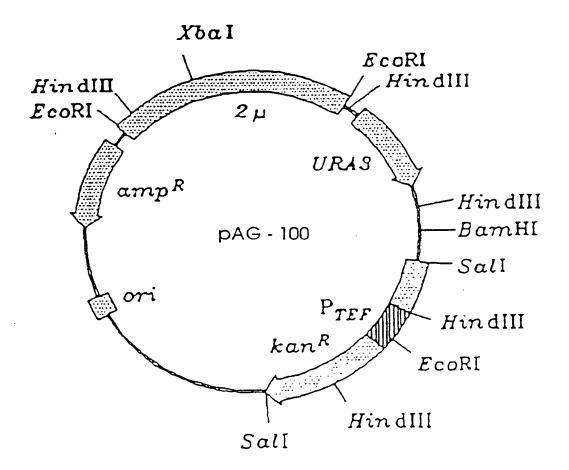
Fraktionen des Sau 3A - Verdaus nach Ultrazentrifugation

Figur 4

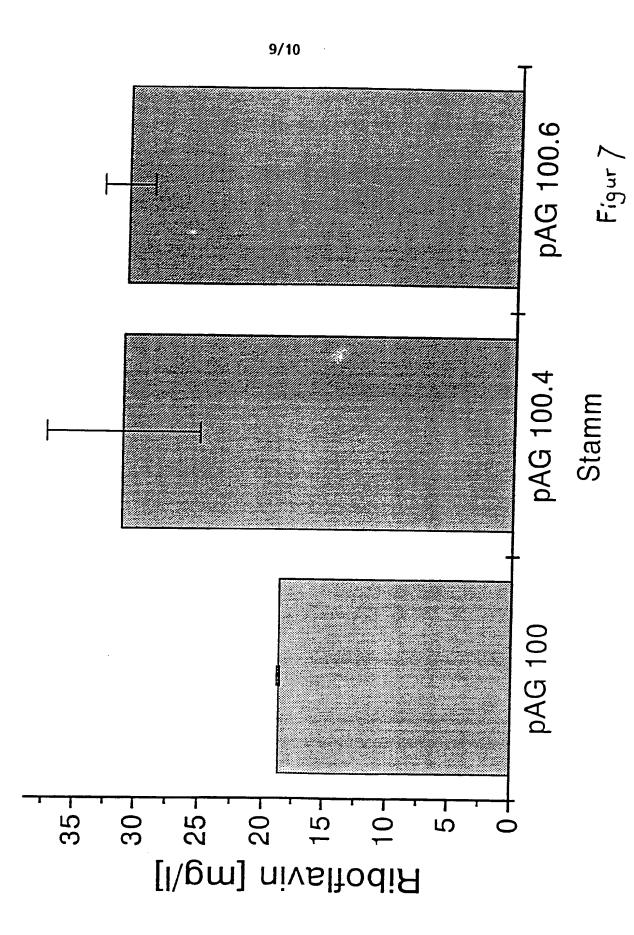


BNSDOCID: <WO\_\_\_9703208A1\_i\_>

8/10



Figur 6



BNSDOCID: <WO\_\_\_9703208A1\_!\_>

**European Patent Office** 

Office européen des brevets



(11) EP 0 927 761 A2

(12)

## **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag: 07.07.1999 Patentblatt 1999/27

(21) Anmeldenummer: 98123331.5

(22) Anmeldetag: 08.12.1998

(51) Int. Cl.<sup>6</sup>: **C12N 15/52**, C12N 15/53, C12N 15/54, C12P 25/00, C12N 9/00, C12N 9/04, C12N 9/10, C12N 9/12

(84) Benannte Vertragsstaaten:

AL LT LV MK RO SI

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE Benannte Erstreckungsstaaten:

(30) Priorität: 23.12.1997 DE 19757755

(71) Anmelder:

BASF AKTIENGESELLSCHAFT 67056 Ludwigshafen (DE)

(72) Erfinder:

Pompejus, Markus Dr.
 67165 Waldsee (DE)

 Seulberger, Harald Dr. 67141 Neuhofen (DE)

 Höffken, Hans Wolfgang Dr. 67069 Ludwigshafen (DE)

 Revuelta Doval, Jose Luis Prof.-Dr. 37001 Salamanca (ES)

 Jimenez, Alberto 37006 Salamanca (ES)

 Santos Garcia, Maria Angeles Dr. 37009 Salamanca (ES)

(54) Gene der Purinbiosyntese aus Ashbya gossypii und deren Verwendung in der mikrobiellen Riboflavinsynthese

(57) Gene der Purinbiosynthese aus Ashbya gossypii und deren Verwendung in der mikrobiellen Riboflavinsynthese.

## Beschreibung

[0001] Die vorliegende Erfindung betrifft Gene der Purinbiosynthese aus Ashbya gossypii und deren Verwendung in der Riboflavinsynthese.

[0002] Vitamin B2, auch Riboflavin genannt, ist für Mensch und Tier essentiell. Bei Vitamin B2-Mangel treten Entzündungen der Mund- und Rachenschleimhäute, Juckreiz und Entzündungen in den Hautfalten und ähnliche Hautschäden, Bindehautentzündungen, verminderte Sehschärfe und Trübung der Hornhaut auf. Bei Säuglingen und Kindern können Wachstumsstillstand und Gewichtsabnahme eintreten. Vitamin B2 hat daher wirtschaftliche Bedeutung insbesondere als Vitaminzusatz bei Vitaminmangel und als Futtermittelzusatz. Daneben wird es als Lebensmittelfarbstoff, beispielsweise in Mayonnaise, Eiscreme, Pudding etc. eingesetzt.

[0003] Die Herstellung von Vitamin B2 erfolgt entweder chemisch oder mikrobiell (siehe z.B. Kurth et al. (1996) Riboflavin, in: Ullmann's Encyclopedia of industrial chemistry, VCH Weinheim). Bei den chemischen Herstellverfahren wird Riboflavin in der Regel in mehrstufigen Prozessen als reines Endprodukt gewonnen, wobei relativ kostspielige Ausgangsprodukte, wie z.B. D-Ribose, eingesetzt werden müssen. Eine Alternative zur chemischen Synthese von Riboflavin ist die Herstelltung dieses Stoffes durch Mikroorganismen. Als Ausgangsstoffe dienen dabei nachwachsende Rohstoffe, wie Zucker oder pflanzliche Öle. Die Herstellung von Riboflavin durch Fermentation von Pilzen wie Eremothecium ashbyii oder Ashbya gossypii ist bekannt (The Merck Index, Windholz et al., eds. Merck & Co., Seite 1183, 1983), aber auch Hefen, wie z.B. Candida, Pichia und Saccharomyces oder Bakterien, wie z.B. Bacillus, Clostridien oder Corynebakterien sind als Riboflavin-Produzenten beschrieben.

[0004] In der EP 405370 sind Riboflavin-überproduzierende Bakerienstämme beschrieben, die durch Transformation der Riboflavin-Biosynthese-Gene aus Bacillus subtilis erhalten wurden. Diese dort beschriebenen Gene und andere, an der Vitamin B2-Biosynthese beteiligten Gene aus Prokaryonten sind für ein rekombinantes Riboflavin-Herstellverfahren mit Eukaryonten, wie z.B. Saccharomyces cerevisiae oder Ashbya gossypii ungeeignet.

[0005] In der DE 44 20 785 sind sechs Riboflavin-Biosynthesegene aus Ashbya gossypii beschrieben, sowie Mikroorganismen, die mit diesen Genen transformiert wurden und die Verwendung solcher Mikroorganismen zur Riboflavinsynthese.

[0006] Mit diesen Verfahren ist es möglich, Produktionsstämme für die mikrobielle Riboflavinsynthese zu erzeugen. Diese Produktionsstämme besitzen jedoch häufig Stoffwechsellimitierungen, die durch die insertierten Biosynthesegene nicht beseitigt werden können oder manchmal erst dadurch entstehen. Solche Produktionsstämmen können manchmal nicht genügend Substrat zur Sättigung mancher Biosyntheseschritte liefern, so daß die Biosynthesekapazität mancher Stoffwechselabschnitte gar nicht voll ausgeschöpft werden kann.

[0007] Daher ist es wünschenswert, weitere Teilbereiche von Stoffwechselwegen zu verstärken, dadurch Stoffwechselengpässe zu beseitigen und dadurch dann die für die mikrobielle Riboflavinsynthese eingesetzten Mikroorganismen (Produktionsstämme) bezüglich ihrer Fähigkeit zur Riboflavinsynthese weiter zu optimieren. Es ist anzustreben, die zu verstärkenden Teilbereiche des komplexen Stoffwechsels zu identifizieren und diese auf geeignete Weise zu verstärken.

[0008] Die Erfindung betrifft neue Proteine der Purinbiosynthese, deren Gene und deren Verwendung für die mikrobielle Riboflavinsynthese.

[0009] Der Purinstoffwechsel (für eine Übersicht siehe z.B. Voet, D. und Voet, J.G., 1994, Biochemie, VCH Weinheim, Seite 743-771; Zalkin, H. und Dixon, J.E., 1992, De novo purine nucleotide biosynthesis, in: Progress in nucleic acid research and molecular biology, Vol. 42, Seite 259-287, Academic Press) ist ein für alle Lebewesen essentieller Teil des Stoffwechsels. Fehlerhafter Purinstoffwechsel kann beim Menschen zu schweren Krankheiten führen (z.B. Gicht). Der Purinstoffwechsel ist zudem ein wichtiges target für die Therapie von Tumorerkrankungen und Virusinfektionen. Zahllose Publikationen sind erschienen, die Substanzen für diese Indikationen beschreiben die im Purinstoffwechsel eingreifen (als Übersicht z.B. Christopherson, R.I. und Lyons, S.D., 1990, Potent inhibitors of de novo pyrimidine and purine biosynthesis as chemotherapeutic agents, Med. Res. Reviews 10, Seite 505-548).

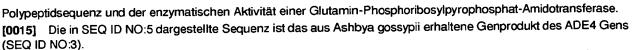
[0010] Untersuchungen der in Purinstoffwechsel beteiligten Enzyme (Smith, J.L., Enzymes in nucleotide synthesis, 1995, Curr. Opinion Struct. Biol. 5, 752-757) zielen darauf ab, neue immunsuppressiv, anti-parasitär oder anti-proliferierend wirkende Medikamente zu entwickeln (Biochem. Soc. Transact. 23, Seite 877-902, 1995).

[0011] Bei diesen Medikamenten handelt es sich dann üblicherweise um natürlich nicht vorkommende Purine, Pyrimidine oder davon abgeleiteter Verbindungen.

[0012] Gegenstand der Erfindung ist ein Protein mit der in SEQ ID NO:2 dargestellten Polypeptidsegenz oder einer aus SEQ ID NO:2 durch Substitution, Insertion oder Deletion von bis zu 15% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer Phosphoribosylpyrophosphat-Synthetase.

55 [0013] Die in SEQ ID NO:2 dargestellte Sequenz ist das aus Ashbya gossypii erhaltene Genprodukt des KPR1 Gens (SEQ ID NO:1).

[0014] Ein weiterer Gegenstand der Erfindung ist ein Protein mit der in SEQ ID NO:5 dargestellten Polypeptidseqenz oder einer aus SEQ ID NO:5 durch Substitution, Insertion oder Deletion von bis zu 10% der Aminosäuren erhältlichen



[0016] Ein weiterer Gegenstand der Erfindung ist ein Protein mit der in SEQ ID NO:8 dargestellten Polypeptidseqenz oder einer aus SEQ ID NO:8 durch Substitution, Insertion oder Deletion von bis zu 20% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer IMP-Dehydrogenase.

[0017] Die in SEQ ID NO:8 und 9 dargestellte Sequenz ist das aus Ashbya gossypii erhaltene Genprodukt des GUA1 Gens (SEQ ID NO:7).

[0018] Ein weiterer Gegenstand der Erfindung ist ein Protein mit der in SEQ ID NO:11 dargestellten Polypeptidseqenz oder einer aus SEQ ID NO:11 durch Substitution, Insertion oder Deletion von bis zu 10% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer GMP-Synthetase.

[0019] Die in SEQ ID NO:11 dargestellte Sequenz ist das aus Ashbya gossypii erhaltene Genprodukt des GUA2 Gens (SEQ ID NO:10).

[0020] Ein weiterer Gegenstand der Erfindung ist ein Protein mit der in SEQ ID NO:13 dargestellten Polypeptidseqenz oder einer aus SEQ ID NO:13 durch Substitution, Insertion oder Deletion von bis zu 10% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer Phosphoribosylpyrophosphat-Synthetase.

[0021] Die in SEQ ID NO:13 dargestellte Sequenz ist das aus Ashbya gossypii erhaltene Genprodukt des KPR2 Gens (SEQ ID NO:12).

[0022] Diese genannten Genprodukte können durch übliche Verfahren der Gentechnologie wie ortsgerichtete Mutagenese so verändert werden, daß bestimmte Aminosäuren ausgetauscht, zusätzlich eingeführt oder entfernt werden. Üblicherweise (aber nicht ausschließlich) werden Aminosäurereste durch solche ähnlicher Raumerfüllung, Ladung oder Hydrophilie/Hydrophobie ausgetauscht um die enzymatischen Eigenschaften der Genprodukte nicht zu verlieren. Insbesondere im aktiven Zentrum führen Veränderungen der Aminosäuresequenz oft zu drastischer Veränderung der enzymatischen Aktivitäten. An anderen, weniger essentiellen Stellen werden jedoch Veränderungen der Aminosäuresequenz häufig toleriert.

[0023] Bei den erfindungsgemäßen Proteinen können

- 1. im Fall des Genproduktes des AgKPR1-Gens, bis zu 15, bevorzugt bis zu 10 und besonders bevorzugt bis zu 5% der Aminosäuren gegenüber der im Sequenzprotokoll dargestellten Sequenzen verändert sein;
- 2. im Fall des Genproduktes des AgADE4-Gens, bis zu 10 und besonders bevorzugt bis zu 5% der Aminosäuren gegenüber der im Sequenzprotokoll dargestellten Sequenzen verändert sein;
- 3. im Fall des Genproduktes des AgGUA1-Gens, bis zu 20, bevorzugt bis zu 15, besonders bevorzugt bis zu 10 und insbesondere bevorzugt bis zu 5% der Aminosäuren gegenüber der im Sequenzprotokoll dargestellten Sequenzen verändert sein;
- 4. im Fall des Genproduktes des AgGUA2-Gens, bis zu 10 und besonders bevorzugt bis zu 5% der Aminosäuren gegenüber der im Sequenzprotokoll dargestellten Sequenzen verändert sein;
- 5. im Fall des Genproduktes des AgKPR2 Gens bis zu 10 %, bevorzugt bis zu 7 % und besonders bevorzugt bis zu 5 % der Aminosäuren gegenüber der im Sequenzprotokoll dargestellten Sequenzen verändert sein.

[0024] Bevorzugt sind solche Proteine, die zwar noch die jeweilige enzymatische Aktivität besitzen, aber in ihrer Regulation verändert worden sind. Viele dieser Enzyme unterliegen einer starken Aktivitätskontrolle durch Zwischen- und Endprodukte (feedback-Inhibition). Dies führt dazu, daß die Enzyme, sobald genügend Endprodukt vorhanden ist, in ihrer Aktivität gedrosselt werden.

[0025] Diese im physiologischen Zustand ökonomische Regelung führt bei Produktionsstämmen jedoch häufig dazu, daß die Produktivität nicht über eine gewisse Grenze hinaus gesteigert werden kann. Durch Beseitigung einer solchen feedback-Inhibition erreicht man, daß die Enzyme ungeachtet der Endproduktkonzentration ihre Aktivität beibehalten und dadurch Stoffwechselengpässe umgangen werden. Dies führt letztlich zu einer deutlichen Steigerung der Riboflavinhinsverthese

[0026] Bevorzugte erfindungsgemäße Proteine sind solche, die nicht mehr durch Folgeprodukte von Stoffwechselwegen (die von Produkten der Enzyme ausgehen) gehemmt werden. Besonders bevorzugte erfindungsgemäße Proteine sind solche, die nicht mehr durch Zwischenprodukte der Purinbiosynthese, insbesondere durch Purinbasen, Purinnukleoside, Purinnukleotid-5'-Monophosphate oder Purinnukleotid-5'-Diphosphate oder Purinnukleotid-5'-Triphosphate gehemmt werden. Insbesondere bevorzugte erfindungsgemäße Proteine sind solche, mit nachfolgenden Veränderungen der Aminosäuresequenz und alle Kombinationen von Aminosäuresequenz-Veränderungen, die diese nachfolgen-

30

35

den Veränderungen enthalten.

[0027] Veränderungen der Aminosäuresequenz am AgKPR1 Genprodukt:

Lysin an Position 7 ausgetauscht gegen Valin Aspartat an Position 52 ausgetauscht gegen Histidin Leucin an Position 131 ausgetauscht gegen Isoleucin Aspartat an Position 186 ausgetauscht gegen Histidin Alanin an Position 193 ausgetauscht gegen Valin Histidin an Position 196 ausgetauscht gegen Glutamin

[0028] Veränderungen der Aminosäuresequenz am AgADE4 Genprodukt:

Aspartat an Position 310 ausgetauscht gegen Valin Lysin an Position 333 ausgetauscht gegen Alanin Alanin an Position 417 ausgetauscht gegen Tryptophan

[0029] Die folgenden Beispiele beschreiben die Herstellung der erfindungsgemäßen Proteine und Nukleinsäuren sowie deren Verwendung zur Herstellung von Mikroorganismen mit gesteigerter Riboflavinsynthese.

Beispiel 1:

5

10

15

25

Herstellung einer genomischen Genbank aus Ashbya gossypii ATCC10895

[0030] Genomische DNA aus Ashbya gossypii ATCC10895 kann nach üblichen Verfahren präpariert werden, z.B. wie beschrieben in EP9703208. Die genomische Genbank, ausgehend von dieser DNA, kann nach üblichen Methoden (z.B. Sambrook, J. et al. (1989) Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press oder Ausubel, F.M. et al. (1994) Current protocols in molecular biology, John Wiley and sons) in beliebigen Plasmiden oder Cosmiden, wie z.B. SuperCos1 (Stratagene, La Jolla, USA) erstellt werden.

Beispiel 2:

Klonierung des Gens für PRPP-Synthetase aus Ashbya gossypii ATCC10895 (AgKPR1)

[0031] Die Klonierung des Gens für PRPP-Synthetase aus Ashbya gossypii (AgKPR1) kann über zwei Schritte verlaufen. Im ersten Schritt kann mit folgenden Oligonukleotiden einen definierten Bereich des KPR1 Gens aus genomischer DNA von Ashbya gossypii über PCR amplifizieren:

KPR5: 5'- GATGCTAGAGACCGCGGGGTGCAAC -3'

KPR3: 5'- TGTCCGCCATGTCGTCTACAATAATA -3'

[0032] Die PCR kann nach üblicher Methode durchgeführt werden. Das resultierende 330 bp DNA Fragment kann mit üblichen Methoden in den Vektor pGEMT (Promega, Madison, USA) kloniert und sequenziert werden.

[0033] Mit dieser Nukleotidsequenz als Sonde kann mit üblichen Methoden eine genomische Cosmid-Genbank gescreent werden. Von einem Cosmid, das ein Signal mit dieser Sonde ergibt, kann dann ein 1911 bp Pstl-HindIII Fragment in den Vektor pBluescript SK+ (Stratagene, La Jolla, USA) subkloniert werden. Auf diesem Fragment liegen das KPR1 Gen und unvollständige ORFs, die Homologie zeigen zu den UBC6 und UBP9 Genen aus Saccharomyces cerevisiae.

[0034] Die PRPP-Synthetase KPR2 und die putative PRPP-Synthetase KPR4 aus Saccharomyces cerevisiae sind die Enzyme, die der PRPP-Synthetase aus Ashbya gossypii mit Ähnlichkeiten von 80,2% bzw. 79,6% am verwandtesten sind. Die KPR2 und KPR4 Gene aus Saccharomyces cerevisiae sind zu 67.6% bzw. 67.8% ähnlich zum KPR1 Gen aus Ashbya gossypii. Andere Enzyme bzw. Gene aus anderen Organismen sind deutlich verschiedener zum KPR1 Gen bzw. zur PRPP-Synthetase aus Ashbya gossypii.

[0035] Die Sequenzvergleiche können z.B. mit dem Clustal Algorithmus mit Hilfe der PAM250 Gewichtungstabelle oder dem Wilbur-Lipman DNA alignment Algorithmus (wie z.B. in dem Programmpaket MegAlign 3.06 der Firma DNA-star implementiert) durchgeführt werden. Mit dem beschriebenen Oligonukleotide-Paar ist es nicht möglich, die Gene für die verschiedenen PRPP-Synthetasen aus Saccharomyces cerevisiae zu amplifizieren.

[0036] Mit der Sonde kann man auch noch einen Klon aus der Genbank finden. Dieser zweite Klon zeigte ein Gen, das ebenfalls für eine PRPP Synthetase kodiert. Dieses Gen wird AgKPR2 genannt und ist deutlich verschieden zu AgKPR1. AgKPR2 zeigt auf Aminosäureebene 66 % Identität zu AgKPR1. Das AgKPR2 Gen (SEQ ID NO: 12) wurde mit allen Proteinen der Swissprot Datenbank verglichen. Die maximale Ähnlichkeit zeigt dieses Protein (88 % Identität bzw. 95 % Äknlichkeit) zum KPR3 Genprodukt aus Saccharomyces cerevisiae. Das Genprodukt des AgKPR1 Gens ist für den überwiegenden Teil der Aktivität der PRPP Synthetase bei Ashbya gossypii verantwortlich. Wenn man das AgKPR1 Gen von Ashbya gossypii disruptiert (analog zur Disruption anderer Ashbya Gene, wie in den Beschreibungen in den Beispielen 6-8), dann findet man deutlich verringerte Enzymaktivität: statt 22 U/mg Protein nur noch 3 U/mg Pro-

tein. Zur Analytik siehe Beispiel 13. In Beispiel 11, 13 und 15 werden Ausführungsbeispiele mit dem AgKPR1 Gen gezeigt, man kann solche Arbeiten aber auch mit AgKPR2 durchführen.

Beispiel 3:

5

Klonierung des Gens für Glutamin-PRPP-Amidotransferase aus Ashbya gossypii ATCC10895 (AgADE4)

[0037] Die Klonierung des Gens für Glutamin-PRPP-Amidotransferase aus Ashbya gossypii (AgADE4) kann über zwei Schritte verlaufen. Im ersten Schritt kann mit folgenden Oligonukleotiden einen definierten Bereich des AgADE4 Gens aus genomischer DNA von Ashbya gossypii über PCR amplifizieren:

ADE4A: 5'- ATATCTTGATGAAGACGTTCACCGT -3'

ADE4B: 5'- GATAATGACGGCTTGGCCGGGAAGA -3'

[0038] Die PCR kann nach üblicher Methode durchgeführt werden. Das resultierende 360 bp DNA Fragment kann mit üblichen Methoden in den Vektor pGEMT (Promega, Madison, USA) kloniert und dann sequenziert werden.

[0039] Mit dieser Sequenz als Sonde kann mit üblichen Methoden eine genomische Cosmid-Genbank gescreent werden. Von einem Cosmid, das ein Signal mit dieser Sonde ergibt, kann dann ein 5369 bp Hindlll Fragment in den Vektor pBluescript SK+ (Stratagene, La Jolla, USA) subkloniert werden. Auf diesem Fragment liegen das AgADE4 Gen und das Gen für das Ashbya Homolog für den mitochondrialen ABC Transporter ATM1 aus Saccharomyces cerevisiae und ein weiterer offener Leseraster, dessen Funktion nicht bekannt ist.

[0040] Das AgADE4 Genprodukt (Glutamin-PRPP-Amidotransferase) zeigt die deutlichste Ähnlichkeit zu den ADE4 Genprodukten aus Saccharomyces cerevisiae und Saccharomyces kluyveri (81% bzw. 86.3%). Die entsprechenden Gene sind jedoch nur zu 68.8% bzw. 72% homolog. Die Ähnlichkeit zu anderen Glutamin-PRPP-Amidotransferasen ist deutlich geringer (z.B. nur 27.5% Ähnlichkeit zum entsprechenden Enzym aus Bacillus subtilis). Die Sequenzvergleiche kann man so durchführen, wie in Beispiel 2 beschrieben.

[0041] Es ist mit dem beschriebenen Paar an Oligonukleotiden nicht möglich, die ADE4 Gene aus Saccharomyces cerevisiae oder Saccharomyces kluyveri zu amplifizieren.

Beispiel 4:

Klonierung des Gens für Inosin-Monophosphat-Dehydrogenase aus Ashbya gossypii ATCC10895 (AgGUA1)

[0042] Die Klonierung des Gens für Inosin-Monophosphat-Dehydrogenase aus Ashbya gossypii (AgGUA1) kann über zwei Schritte verlaufen.

[0043] Im ersten Schritt kann man mit folgenden Oligonukleotiden einen definierten Bereich des AgGUA1 Gens aus genomischer DNA von Ashbya gossypii über PCR amplifizieren:

IMP5: 5'- GGCATCAACCTCGAGGAGGCGAACC -3'

IMP3: 5'- CAGACCGGCCTCGACCAGCATCGCC - 3'

[0044] Die PCR kann nach üblicher Methode durchgeführt werden. Das resultierende 230 bp DNA Fragment kann mit üblichen Methoden in den Vektor pGEMT (Promega, Madison, USA) kloniert und dann sequenziert werden.

[0045] Mit dieser Sequenz als Sonde kann mit üblichen Methoden eine genomische Cosmid-Genbank gescreent werden. Von einem Cosmid, das ein Signal mit dieser Sonde ergibt, kann ein 3616 bp Apal Fragment in den Vektor pBluescript SK+ (Stratagene, La Jolla, USA) subkloniert werden. Die kodierende Region des AgGUA1 Gens Gens ist 1569 bp lang und ist unterbrochen von einem161 bp langen Intron. Die Intron Grenzen (5' splice site AGGTATGT und 3' splice site CAG) kann man durch Klonierung und Sequenzierung von AgGUA1cDNA verifizieren.

[0046] AgGUA1 ist das erste beschriebene Gen aus Ashbya gossypii mit einem Intron.

[0047] Das AgGUA1 Genprodukt (IMP-Dehydrogenase) zeigt die deutlichste Ähnlichkeit zu den 4 IMP-Dehydrogenasen aus Saccharomyces cerevisiae (Ähnlichkeiten zwischen 67% und 77.2%). Die Ähnlichkeit zu anderen IMP-Dehydrogenasen ist deutlich geringer Die Sequenzvergleiche kann man so durchführen, wie in Beispiel 2 beschrieben. Ashbya gossypii scheint nur ein Gen für dieses Enzym zu haben. Dies kann man durch southern blotting mit genomischer DNA von Ashbya gossypii mit Hilfe der oben genannter Sonde zeigen.

[0048] Das Gen aus Saccharomyces cerevisiae, das für die dem AgGUA1 Genprodukt ähnlichste IMP-Dehydrogenase kodiert (IMH3), hat eine Ähnlichkeit von 70.2% zum AgGUA1 Gen. Es ist mit dem beschriebenen Paar an Oligonukleotiden nicht möglich, dieses Gen aus Saccharomyces cerevisiae zu amplifizieren.

#### Beispiel 5:

Klonierung des Gens für Guanosin-Monophosphat-Svnthetase aus Ashbya gossypii ATCC10895 (AgGUA2)

[0049] Die Klonierung des Gens für Guanosin-Monophosphat-Synthetase aus Ashbya gossypii (AgGUA2) kann über zwei Schritte verlaufen. Im ersten Schritt kann man mit folgenden Oligonukleotiden einen definierten Bereich des AgGUA2 Gens aus genomischer DNA von Ashbya gossypii über PCR amplifizieren:

GUA2A: 5'- TGGACCGGGCGGTGTTCGAGTTGGG -3'

GUA2B: 5'- AGGCTGGATCCTGGCTGCCTCGCGC -3'

[0050] Die PCR Reaktion kann nach üblicher Methode durchgeführt werden. Das resultierende 750 bp DNA Fragment kann mit üblichen Methoden in den Vektor pBluescript SK+ (Stratagene, La Jolla, USA) kloniert und dann sequenziert werden.

[0051] Mit dieser Sequenz als Sonde kann mit üblichen Methoden eine genomische Cosmid-Genbank gescreent werden. Von einem Cosmid, das ein Signal mit dieser Sonde ergibt, kann dann ein 2697 bp Clal-EcoRV Fragment in den Vektor pBluescript SK+ (Stratagene, La Jolla, USA) subkloniert werden.

[0052] Das AgGUA2 Genprodukt (GMP-Synthetase) zeigt die deutlichste Ähnlichkeit zu GMP-Synthetase aus Saccharomyces cerevisiae (Ähnlichkeiten 86.6%). Die Gene für die GMP-Synthetasen aus Saccharomyces cerevisiae und Ashbya gossypii sind zu 71.2% homolog. Die Ähnlichkeit des AgGUA2 Genproduktes zu anderen IMP-Dehydrogenasen ist deutlich geringer Die Sequenzvergleiche kann man so durchführen, wie in Beispiel 2 beschrieben.

[0053] Es ist mit dem beschriebenen Paar an Oligonukleotiden nicht m\u00f6glich, das GMP-Synthetase Gen aus Saccharomyces cerevisiae zu amplifizieren.

#### Beispiel 6:

25 Disruption des AgADE4 Gens von Ashbya gossypii ATCC10895

[0054] Unter Disruption eines Genes versteht man die Zerstörung der Funktionalität einer genomischen Kopie des Gens entweder durch (a) Entfernen eines Teiles der Gensequenz, oder durch (b) der Unterbrechung des Gens durch Einfügung eines Stückes Fremd-DNA in das Gen oder durch (c) Ersatz eines Teil des Gens durch Fremd-DNA. Die verwendete Fremd-DNA ist beliebig, bevorzugt aber ein Gen, das Resistenz gegen eine beliebige Chemikalie bewirkt. Zur Disruption von Genen können beliebige Resistenzgene verwendet werden.

[0055] Zur Disruption des AgADE4-Gens von Ashbya gossypii ATCC10895 kann man ein Gen verwenden, das Resistenz gegen G418 vermittelt. Es kann sich dabei um das Kanamycin-Resistenzgen aus TN903, unter Kontrolle des TEF-Promotors von Ashbya gossypii (siehe z.B. Yeast 10, S.1793-1808, 1994, WO9200379) handeln. Das Gen ist 5' und 3' von mehreren Schnittstellen für Restriktionsendonukleasen flankiert, so daß eine Kassette aufgebaut wurde, die beliebige Konstruktionen von Gen-Disruptionen mit üblichen Methoden der in vitro Manipulation von DNA ermöglichen. [0056] Das interne HincII Fragment von AgADE4 (zwischen den Positionen 2366 und 2924) kann durch eine wie oben skizzierte Resistenzkassette ersetzt werden. Das erhaltene Konstrukt erhält den Namen ade4::G418.

[0057] Das erhaltene Plasmid kann man in E.coli vermehren. Das BamHI / BgIII- Fragment des Konstruktes ade4::G418 kann präpariert, über Agarosegel-Elektrophorese und nachfolgende Elution der DNA aus dem Gel (siehe Proc. Natl. Acad. Sci. USA 76, 615-619, 1979) aufgereinigt und zur Transformation von Ashbya gossypii eingesetzt werden.

[0058] Ashbya gossypii kann durch Protoplastentransformation (Gene 109, 99-105, 1991), bevorzugt aber durch Elektroporation (BioRad Gene Pulser, Bedingungen: Küvetten mit Spaltbreite 0,4 mm, 1500V,  $25\mu$ F,  $100\Omega$ ) transformiert werden. Die Selektion transformierter Zellen erfolgt auf G418-haltigem Festmedium.

[0059] Erhaltene G418-resistente Klone können mit üblichen Methoden der PCR und Southern-Blot Analyse daraufhin untersucht werden, ob die genomische Kopie des AgADE4 Gens tatsächlich zerstört ist. Klone, deren AgADE4 Gen zerstört ist, sind Purin- auxotroph.

### 50 Beispiel 7:

Disruption des AgGUA1 Gens von Ashbya gossypii ATCC10895

[0060] Zur prinzipiellen Beschreibung der Disruption von Genen, der Verwendung einer Resistenzkassette und der Transformation von Ashbya gossypii siehe Beispiel 6.

[0061] Das interne Xhol / Kpnl Fragment von AgGUA1 (zwischen den Positionen 1620 und 2061) kann durch eine wie oben skizzierte Resistenzkassette ersetzt werden. Das erhaltene Konstrukt erhält den Namen gua1::G418.

[0062] Das erhaltene Plasmid kann in E.coli vermehrt werden. Das Xbal / BamHI - Fragment des Konstruktes



gua1::G418 kann präpariert, über Agarosegel-Elektrophorese und nachfolgende Elution der DNA aus dem Gel aufgereinigt und zur Transformation von Ashbya gossypii eingesetzt werden.

[0063] Erhaltene G418-resistente Klone können mit üblichen Methoden der PCR und Southern-Blot Analyse daraufhin untersucht werden, ob die genomische Kopie des AgGUA1 Gens tatsächlich zerstört ist. Klone, deren AgGUA1 Gen zerstört ist, sind Guanin- auxotroph.

### Beispiel 8:

10

25

40

55

Disruption des AgGUA2 Gens von Ashbya gossypii ATCC10895

[0064] Zur prinzipiellen Beschreibung der Disruption von Genen, der Verwendung einer Resistenzkassette und der Transformation von Ashbya gossypii siehe Beispiel 6.

[0065] Das interne Sall Fragment von AgGUA2 (zwischen den Positionen 1153 und 1219) kann man durch eine wie oben skizzierte Resistenzkassette ersetzen. Das erhaltene Konstrukt erhält den Namen gua2::G418.

[0066] Das erhaltene Plasmid kann in E.coli vermehrt werden. Das Xbal / BamHl - Fragment des Konstruktes gua2::G418 kann präpariert, über Agarosegel-Elektrophorese und nachfolgende Elution der DNA aus dem Gel aufgereinigt und zur Transformation von Ashbya gossypii eingesetzt werden.

[0067] Erhaltene G418-resistente Klone können mit üblichen Methoden der PCR und Southern-Blot Analyse daraufhin untersucht werden, ob die genomische Kopie des AgGUA2 Gens tatsächlich zerstört ist. Klone, deren AgGUA2 Gen zerstört ist, sind Guanin- auxotroph.

#### Beispiel 9:

Klonierung des GAP-Promotors aus Ashbya gossypii

[0068] Das Gen für Glycerinaldehyd-3-Phosphat-Dehydrogenase aus Ashbya gossypii (AgGAP) kann man durch ein allgemein übliches Screening einer genomischen Ashbya gossypii Cosmid-Genbank (siehe Beispiel 1, mit einer Sonde, die aus Sequenzinformationen des GAP Gens aus Saccharomyces cerevisiae erstellt wurde) klonieren.

[0069] Der 5' nicht-translatierte Bereich des Gens (-373 bis -8 Region, bezogen auf den Translationsstart) wurde als Promotor angenommen. Flankierend zu dieser Sequenz wurden 2 Schnittstellen für die Restriktionsendonuklease Not1 eingeführt. In diesem Bereich findet man die bona fide TATA Box (nt 224-230), zwei Sequenzabschnitte (nt 43-51 und 77-85), die dem sogenannten GCR1 binding element und einen Sequenzabschnitt, (nt 9-20) dessen Komplement partial dem RAP1 binding element von Saccharomyces cerevisiae entspricht (siehe z.B. Johnston, M. und Carlson, M. (1992) pp.193-281 in The molecular biology and cellular biology of the yeast Saccharomyces: Gene expression, Cold Spring Harbor Laboratory Press). Die so konstruierte Promotorkassette kann als einfach portierbares Expressionsignal vor jedes beliebige Gen für die Überexpression in Ashbya gossypii gesetzt werden und führt zu deutlicher Überexpression von Genen in Ashbya gossypii, wie gezeigt in Beispiel 11.

### Beispiel 10:

Konstruktion von Plasmiden mit Genen unter Kontrolle des GAP-Promotors aus Ashbya gossypii

[0070] Zur Einfügung der GAP-Promotorkassette 5i der kodierenden Region des AgADE4 Gens wurde nach üblichen Methode (z.B. Clover, D.M. und Hames, B.D. (1995) DNA cloning Vol.1, IRL press) 8 bp 5' des ATG Startcodons eine singuläre Notl Schnittstelle (Erkennungssequenz GCGGCCGC) eingeführt.

[0071] Die GAP-Promotorkassette kann dann über Notl in diese Position eingefügt werden. Analog kann man vorgehen bei der Klonierung der GAP-Promotorkassette 5' der kodierenden Region der Gene AgKPR1, AgGUA1, AgGUA2 sowie bei Varianten der Gene AgADE4, AgKPR1, AgGUA1 und AgGUA2.

[0072] In Ashbya gossypii wird die Expression der Gene, die die GAP-Promotorkassette 5' der kodierenden Region tragen durch den GAP-Promotor kontrolliert.

### Beispiel 11:

Überexpression Genen in Ashbya gossypii unter Kontrolle des GAP-Promotors

[0073] Die Transformation von Ashbya gossypii mit den in Beispiel 10 beschriebenen DNA-Konstrukten kann man durchführen wie in Beispiel 6 beschrieben. Als Empfänger können bevorzugt, aber nicht ausschließlich, solche Klone dienen, die vor der hier durchzuführenden Transformation eine Disruption des zu überexprimierenden Gens tragen. So

kann man z.B. die in Beispiel 6 beschriebene Mutante von Ashbya gossypii, die eine ade4::G418 Mutation trägt, mit einem in Beispiel 10 beschriebenen GAP-ADE4 Konstrukt transformieren. Man kann die Integration des Konstruktes in das Genom durch Southern-Blot-Analyse verifizieren. Die resultierenden Klone tragen kein G418 Resistenzgen mehr (sind somit G418 sensitiv) und sind Purin-prototroph. Die Überexpression kann durch Northern-Blot-Analyse oder Nachweis der enzymatischen Aktivität (wie in Beispiel 12 beschrieben) nachgewiesen werden. Bei Expression des AgADE4 Gens unter dem natürlichen Promotor kann man 0,007 U/mg Protein nachweisen. Bei Expression des AgADE4 Gens unter dem GAP Promotor kann man 0,382 U/mg Protein nachweisen

[0074] Ein Sequenzabschnitt der kodierenden Region des AgADE4 Gens kann als Sonde verwendet werden. Analog kann man mit AgKPR1, AgGUA1, AgGUA2 sowie bei Varianten all dieser Gene vorgehen. Außerdem kann man Kombinationen aus einem dieser Gene, zusammen mit anderen Genen auf diese Weise in das Genom von Ashbya gossypii einbringen.

[0075] Der Ashbya gossypii Wild-Typ hat eine spezifische PRPP Synthetase Aktivität von 22 U/mg Protein (zur Analytik der PRPP-Synthetase siehe Beispiel 13). Bei Expression des AgKPR1-Gens unter dem GAP-Promotor kann man 855 U/mg Protein nachweisen.

#### Beispiel 12:

15

[0076] Varianten des AgADE4 Genproduktes (Glutamin-PRPP-Amidotransferase), die nicht mehr feedback durch Purine oder Zwischenprodukte der Purinsynthese gehemmt werden.

[0077] Glutamin-PRPP-Amidotransferasen werden durch Purin-Nukleotide feedback inhibiert. Diese Inhibition findet man in zahlreichen Organismen (siehe z.B. Switzer, R.L. (1989) Regulation of bacterial Glutamine Phosphoribosylpyrophosphate Amidotransferase, in: Allosteric enzymes pp. 129-151, CRC press, Boca Raton).

[0078] Die Glutamin-PRPP-Amidotransferase aus Ashbya gossypii wird ebenfalls durch AMP oder GMP gehemmt (siehe Abbildung ). Die Aktivität der Glutamin Phosphoribosylpyrophosphat Amidotransferase aus Ashbya gossypii kann man messen, wie beschrieben in Messenger und Zalkin (1979) J. Biol. Chem. 254, Seite 3382-3392.

[0079] Man kann veränderte Glutamin Phosphoribosylpyrophosphat Amidotransferasen konstruieren, die nicht mehr durch Purine gehemmt werden. Es ist offensichtlich, daß die Überexpression solch deregulierter Enzyme den Purinstoffwechsel deutlich mehr verstärken als die Überexpression der feedback inhibierten Enzyme. Veränderungen der Sequenz des AgADE4 Gens können nach üblichen Methoden (z.B. Clover, D.M. und Hames, B.D. (1995) DNA cloning Vol.1, IRL press) vorgenommen werden. Man kann z.B. folgende Aminosäuren der Glutamin Phosphoribosylpyrophosphat Amidotransferase ausgetauschen:

[0080] Das Codon, das für Aspartat an der Position 310 kodiert, kann durch ein Codon ersetzt werden, das für Valin kodiert. Das Codon, das für Lysin an der Position 333 kodiert, kann durch ein Codon ersetzt werden, das für Alanin kodiert. Das Codon, das für Alanin an der Position 417 kodiert, kann durch ein Codon ersetzt werden, das für Tryptophan kodiert. Zusätzlich können AgADE4 Gene konstruiert werden, die Kombinationen dieser Austausche tragen.

[0081] Alle Enzyme, die den D310V, den K333A, den A417W oder jede Kombination von Austauschen tragen, die D310V oder K333A enthalten, zeigen verringerte feedback Inhibition durch AMP und GMP (siehe Abbildung ). Dies kann z.B. nach Expression der Enzyme in Ashbya gossypii (siehe Beispiel 11) gezeigt werden.

#### 40 Beispiel 13:

[0082] Varianten des AgKPR1 Genproduktes (PRPP-Synthetase), die nicht mehr feedback durch Purine oder Zwischenprodukte der Purinsynthese gehemmt werden.

[0083] PRPP-Synthetasen werden feedback durch Purine, Pyrimidine und Aminosäuren inhibiert. Diese Inhibition findet man in zahlreichen Organismen (siehe z.B. Gibson, K.J. et al. (1982) J. Biol. Chem. 257, 2391-2396; Tatibana, M. et al. (1995) Adv., Enzyme Regul. 35, 229-249 und darin zitierte Arbeiten).

[0084] In der Forschung der klinischen Medizin sind Fälle erblicher Gicht beschrieben, deren Basis eine verstärkte Purinbiosynthese ist. Die molekulare Ursache hierfür ist eine sogenannte Superaktivität der humanen PRPP Synthetase (siehe z.B. Amer. J. Med. 55 (1973) 232-242; J. Clin. Invest. 96 (1995) 2133-2141; J. Biol. 268 (1993) 26476-26481). Die Basis hierfür kann eine Mutation sein, die dazu führt, daß das Enzym nicht mehr durch Purine feedback inhibitert wird.

Die Aktivität der PRPP Synthetase aus Ashbya gossypii kann man messen wie in Anal. Biochem. 98 (1979) 254-263 oder J. Bacteriol. 174 (1992) 6852-6856 beschrieben. Die spezifische Aktivität (U/mg) wird über die Menge an entstandenem Produkt definiert (nmol/min/mg Protein).

Man kann veränderte PRPP Synthetasen konstruieren, die nicht mehr durch Purine gehemmt werden. Es ist offensichtlich, daß die Überexpression solch deregulierter Enzyme den Purinstoffwechsel deutlich mehr verstärkt als die Überexpression der feedback inhibierten Enzyme. Veränderungen der Sequenz des AgKPR1 Gens können nach üblichen Methoden (z.B. Glover, D.M. und Hames, B.D. (1995) DNA cloning Vol. 1, IRL press) vorgenommen werden. Man kann



z.B. folgende Aminosäuren der PRPP Synthetase austauschen:

Das Codon, das für Leucin an der Position 131 kodiert, kann durch ein Codon ersetzt werden, das für Isoleucin kodiert. Das Codon, das für Histidin an der Position 196 kodiert, kann durch ein Codon ersetzt werden, das für Glutamin kodiert. Alle Enzyme, die einen dieser Aminosäureaustausche (L 131 oder H196Q) tragen, zeigen verringerte feedback Hemmung durch Purine. In Abbildung 2 ist dies gezeigt am Beispiel ADP.

Dies kann gezeigt werden, nachdem die entsprechenden Enzyme in Ashbya gossypii exprimiert wurden. Dies kann entsprechend Beispiel 11 durchgeführt werden.

#### Beispiel 14:

10

[0085] Varianten des AgGUA1 Genproduktes (IMP-Dehydrogenase), die nicht mehr feedback durch Purine oder Zwischenprodukte der Purinsynthese gehemmt werden.

#### Beispiel 15:

15

Auswirkung der Verstärkung und/oder der Optimierung von Purinstoffwechsel- Enzymen und deren Gene auf die Riboflavin-Produktion in Ashbya gossypii

[0086] Man kann den Ausgangsstamm Ashbya gossypii ATCC10895, in Vergleich mit davon abgeleiteten Klonen, die chromosomale Kopien von Genen, unter Kontrolle des GAP Promotors tragen (wie in Beispiel 11 beschrieben), im Schüttelkolben auf Riboflavin- Produktivität prüfen. Man kann dazu 300 ml Schüttelkolben, mit 20 ml YPD Medium (Sambrook, J. et al. (1989) Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press) bei einer Inkubationstemperatur von 28°C einsetzen.

[0087] Nach 2 Tagen produziert der Kontrollstamm durchschnittlich 14,5 mg Riboflavin pro I Kulturbrühe. Stämme, die Gene für Purinstoffwechsel-Enzyme überexprimieren (wie z.B. in Beispiel 11 gezeigt) oder Gene für optimierte Purinstoffwechsel-Enzyme (z.B. wie in den Beispielen 12, 13 ,und 14) überexprimieren, produzieren mehr Riboflavin. So produziert der Stamm, der AgADE4D310VK333A (Beispiel 12) überexprimiert, in 2 Tagen durchschnittlich 45,4 mg Riboflavin pro I Kulturbrühe.

[0088] Der Stamm, der AgKPR1 unter dem GAP-Promotor überexprimiert, produziert statt 14 mg/l (wie der WT) 36 mg/l Riboflavin. Der Stamm, der AgKPR1H196Q unter dem GAP-Promotor überexprimiert, produziert 51 mg/l Riboflavin.

### Abbildung 1:

[0089] Messung der Aktivität der Gln-PRPP-Amidotransferase aus A. gossypii und von veränderten Formen des Enzyms in Abhängigkeit der Konzentration von Adenosin-5'-Monophosphat (AMP) und Guanosin-5'-Monophosphat (GMP).

WT: Gln-PRPP-Amidotransferase

A418W: Gln-PRPP-Amidotransferase, Alanin an Position 418 ausgetauscht gegen Tryptophan.

K333A: Gin-PRPP-Amidotransferase, Lysin an Position 333 ausgetauscht gegen Alanin.

D310VK333A: Gln-PRPP-Amidotransferase, Aspartat an Position 310 ausgetauscht gegen Valin und Lysin an Position 333 ausgetauscht gegen Alanin.

#### 45 Abbildung 2:

[0090] Messung der Aktivität der PRPP Synthetase aus A. gossypii und von veränderten Formen des Enzymes in Abhängigkeit der Konzentration von Adenosin.5'-Diphosphat (ADP)

WT: PRPP Synthetase

L131I: PRPP Synthetase, Leucin an Position 131 ausgetauscht gegen Isoleucin

H196Q: PRPP Synthetase, Histidin an Position 196 ausgetauscht gegen Glutamin

H196Q, L131I: PRPP Synthetase, Histidin an Position 196 ausgetauscht gegen Glutamin und Leucin an Position 131 ausgetauscht gegen Isoleucin

55

50

### SEQUENZPROTOKOLL

	(1) ALGE	MEINE INFORMATION:
5	(i)	ANMELDER:
		(A) NAME: BASF Aktiengesellschaft
		(B) STRASSE: Carl-Bosch-Strasse 38
		(C) ORT: Ludwigshafen
10		(E) LAND: Bundesrepublik Deutschland
		(F) POSTLEITZAHL: D-67056
		(G) TELEPHON: 0621/6048526
		(H) TELEFAX: 0621/6043123
		(I) TELEX: 1762175170
15		
	(11)	ANMELDETITEL: Gene der Purinbiosynthese aus Ashbya gossypii
		und deren Verwendung in der mikrobiellen
		Riboflavinbiosynthese
20	(iii)	ANZAHL DER SEQUENZEN: 13
	(iv)	COMPUTER-LESBARE FORM:
		(A) DATENTRÄGER: Floppy disk
		(B) COMPUTER: IBM PC compatible
25		(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
		(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)
	(2) INFO	RMATION ZU SEQ ID NO: 1:
	(i)	SEQUENZ CHARAKTERISTIKA:
30	(-/	(A) LÄNGE: 1911 Basenpaare
		(B) ART: Nukleinsäure
		(C) STRANGFORM: Einzel
		(D) TOPOLOGIE: linear
35		
	(11)	ART DES MOLEKÜLS: DNS (genomisch)
	(iii)	HYPOTHETISCH: NEIN
	(;;;)	ANTISENSE: NEIN
40	(111)	ANTIOENSE. NEIN
	(ix)	MERKMALE:
		(A) NAME/SCHLÜSSEL: 5'UTR
		(B) LAGE: 1625
45	(ix)	MERKMALE:
		(A) NAME/SCHLÜSSEL: CDS
		(B) LAGE: 6261582
	/4 == \	MERKMALE:
50	(TX)	MERKMALE: (A) NAME/SCHLÜSSEL: 3'UTR
JU		(B) LAGE: 15831911
		(D) DAGE. 13631311

### (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

	GGTAGTCGCT CATCGACAGA CACAATCGCG TGTTCTCTCT GAATCGTCCA TTGGGTGTCA	60
5	GCATCCTGAT CGCGGGCGGA TGGAATGGGT AATCATTAGG AAACACCAAT GTCCCATGGT	120
	ATTGTCCGTC CTCGTATGGT GTCTCAGGAG GACCCGTGAT CACGTAGTGC CACACCAGGA	180
10	TATTGTCTTC CTTTGGTGCT GCCACGATGT AGGGCGGGGG GTTCTCGGTC ATCATTTTGT	240
10	ACTCCTTTGA GAGCCGCTTG TACGCCTGTC TTGATGCCAT CTTGCCTACT ATTAGTTTCT	300
	CACCACTTCC CGCCAAACAA TCTGCACTTT ACGAGCGCTA TCTATCCCTC GGGTCGCTCT	360
15	AGTTGATTAT TGGCGAAACT GATAGTTCAG GTACTTCCAT GATGCGGTCA TATCCACGTA	420
	TGTGATCACG TGATCATCAG CCATGCTGCC AGCTCACGGG CCTGCCTACA CTATTGGAGG	480
	CTCTGTGAGT CATGATTTAT TGCATATCAA GCCCAGATAG TCGTTGGGGA TACTACCGTT	540
20	GCCGCGATGA GCTCCGATAT TAAGTTGTAG CCAAAAATTT TAACGGATGA CTTCTTAACA	600
	GTTATTGACG CCGCAATCCT ACGCC ATG TCG TCC AAT AGC ATA AAG CTG CTA  Met Ser Ser Asn Ser Ile Lys Leu Leu	652
25	1 5	
25	GCA GGT AAC TCG CAC CCG GAC CTA GCT GAG AAG GTC TCC GTT CGC CTA	700
	Ala Gly Asn Ser His Pro Asp Leu Ala Glu Lys Val Ser Val Arg Leu 10 15 20 25	
30	GGT GTA CCA CTT TCG AAG ATT GGA GTG TAT CAC TAC TCT AAC AAA GAG	748
	Gly Val Pro Leu Ser Lys Ile Gly Val Tyr His Tyr Ser Asn Lys Glu 30 35 40	
	ACG TCA GTT ACT ATC GGC GAA AGT ATC CGT GAT GAA GAT GTC TAC ATC	796
35	Thr Ser Val Thr Ile Gly Glu Ser Ile Arg Asp Glu Asp Val Tyr Ile 45 50 55	
	ATC CAG ACA GGA ACG GGG GAG CAG GAA ATC AAC GAC TTC CTC ATG GAA	844
40	Ile Gln Thr Gly Thr Gly Glu Gln Glu Ile Asn Asp Phe Leu Met Glu 60 65 70	
	CTG CTC ATC ATG ATC CAT GCC TGC CGG TCA GCC TCT GCG CGG AAG ATC	892
	Leu Leu Ile Met Ile His Ala Cys Arg Ser Ala Ser Ala Arg Lys Ile	
45	75 80 85	0.40
	ACA GCG GTT ATA CCA AAC TTC CCT TAC GCA AGA CAA GAC AAA AAG GAC Thr Ala Val Ile Pro Asn Phe Pro Tyr Ala Arg Gln Asp Lys Lys Asp	940
	90 95 100 105	
50	AAG TCG CGA GCA CCG ATA ACT GCC AAG CTG GTG GCC AAG ATG CTA GAG	988
	Lys Ser Arg Ala Pro Ile Thr Ala Lys Leu Val Ala Lys Met Leu Glu 110 115 120	

					Asn					Met						CAA Gln	1	1036
5				Phe					Val							CCG Pro	1	084
10			Leu						AAT Asn							ATG Met	1	.132
15		Val							GCG Ala								1	180
20									TTG Leu								1	228
									TTG Leu 210								1:	276
<b>2</b> 5	TCC Ser	TGT Cys	ATT Ile 220	ATT Ile	GTA Val	GAC Asp	GAC Asp	ATG Met 225	GCG Ala	GAC Asp	ACG Thr	TGC Cys	GGA Gly 230	ACG Thr	CTA Leu	GTG Val	1:	324
30									AAT Asn							GCC Ala	13	372
35									GGC Gly								14	420
									AGC Ser								14	468
40									ATT Ile 290								15	516
45	GAG Glu	GCA Ala	ATT Ile 300	AGA Arg	AGG Arg	CTT Leu	His	AAC Asn 305	GGG Gly	GAA Glu	AGT Ser	GTG Val	TCG Ser 310	TAC Tyr	CTG Leu	TTC Phe	15	564
50	AAT Asn	AAC Asn 315	GCT Ala	GTC Val	ATG Met	TAGT	GCTG	TC A	GTGG	CAGA	T GC	ATGA	TCGC	TGG	CCTA	ATT	16	519
	ATCT	GTGT	'AA G	TTGA	TACA	A TG	CAGT	AAAT	' ACA	GTAC	ATA	AAAC	TGAA	TG T	TTTT	CACTT	16	79

	AGGGGTGCTT TGTTGTTCTG ATAGCGTGTG TGCGAATTTG GAGGTGAAAG TTGAACATCA 17	739
	CGTAATGAAT ACAAACAAGA TTGCACATTA GGAAAAGCGA TAAATTATTT ATTATTTGCA 17	799
5	ACTGGCCTTT GAGCGTTTAA GCCTGAACAT TTTTGCCCTT TTGTTTGACC GTACCGTTAT 18	859 <sup>°</sup>
	CACTCGTCCT TATATATGGC TATCCTTCTC TTCCGGAACT TCTTCGAGCG TA	11
10	(2) INFORMATION ZU SEQ ID NO: 2:	
	<ul><li>(i) SEQUENZ CHARAKTERISTIKA:</li><li>(A) LÄNGE: 318 Aminosäuren</li><li>(B) ART: Aminosäure</li></ul>	
4.5	(D) TOPOLOGIE: linear	
15	(ii) ART DES MOLEKÜLS: Protein	
	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:	
20	Met Ser Ser Asn Ser Ile Lys Leu Leu Ala Gly Asn Ser His Pro Asp 1 5 10 15	
	Leu Ala Glu Lys Val Ser Val Arg Leu Gly Val Pro Leu Ser Lys Ile 20 25 30	
25	Gly Val Tyr His Tyr Ser Asn Lys Glu Thr Ser Val Thr Ile Gly Glu 35 40 45	
30	Ser Ile Arg Asp Glu Asp Val Tyr Ile Ile Gln Thr Gly Thr Gly Glu 50 55 60	
	Gln Glu Ile Asn Asp Phe Leu Met Glu Leu Leu Ile Met Ile His Ala 65 70 75 80	
35	Cys Arg Ser Ala Ser Ala Arg Lys Ile Thr Ala Val Ile Pro Asn Phe 85 90 . 95	
	Pro Tyr Ala Arg Gln Asp Lys Lys Asp Lys Ser Arg Ala Pro Ile Thr 100 105 110	
40	Ala Lys Leu Val Ala Lys Met Leu Glu Thr Ala Gly Cys Asn His Val 115 120 125	
	Ile Thr Met Asp Leu His Ala Ser Gln Ile Gln Gly Phe Phe His Ile 130 135 140	
45	Pro Val Asp Asn Leu Tyr Ala Glu Pro Asn Ile Leu His Tyr Ile Gln 145 150 155 160	
50	His Asn Val Asp Phe Gln Asn Ser Met Leu Val Ala Pro Asp Ala Gly 165 170 175	
	Ser Ala Lys Arg Thr Ser Thr Leu Ser Asp Lys Leu Asn Leu Asn Phe 180 185 190	

	Ala	Leu	Ile 195	His	Lys	Glu	Arg	Gln 200	Lys	Ala	Asn	Glu	Val 205	Ser	Arg	Met
5	Val	Leu 210	Val	Gly	Asp	Val	Ala 215	Asp	Lys	Ser	Cys	Ile 220	Ile	Val	Asp	Asp
	Met 225	Ala	Asp	Thr	Cys	Gly 230	Thr	Leu	Val	Lys	A1a 235	Thr	Asp	Thr	Leu	Ile 240
10	Glu	Asn	Cys	Ala	Lys 245	Glu	Val	Ile	Ala	Ile 250	Val	Thr	His	Gly	Ile 255	Phe
15	Ser	Gly	Gly	Ala 260	Arg	Glu	Lys	Leu	Arg 265	Asn	Ser	Lys	Leu	Ala 270	Arg	Ile
	Va1	Ser	Thr 275	Asn	Thr	Val	Pro	Val 280	Asp	Leu	Asn	Leu	Asp 285	Ile	Tyr	His
20	Gln	Ile 290	Asp	Ile	Ser	Ala	Ile 295	Leu	Ala	Glu	Ala	Ile 300	Arg	Arg	Leu	His
	Asn 305	Gly	Glu	Ser	Val	Ser 310	Tyr	Leu	Phe	Asn	Asn 315	Ala	Val	Met		
25	(2)	INFO	RMAT	NOI	ZU S	EQ I	D NO	): 3:								
30		(i)	(A (B (C	) LÄ ) AR ) ST	NGE: T: N RANG	536 lukle FORM	9 Ba insä	nzel	aare	1						
		(ii)	ART	DES	MOL	EKÜL	S: D	NS (	geno	misc	h)					
35				•		H: N NEIN										
40		(ix)	(A		ME/S	СНLÜ 15		: 5′	UTR							
45		(ix)	(A		ME/S		SSEL 1482	: CD	s							
45		(ix)	(A		ME/S		SSEL	: CD	S							
50		(ix)	(A		ME/S		SSEL	: CD	s							

### (ix) MERKMALE:

(A) NAME/SCHLÜSSEL: 3'UTR

(B) LAGE: 4704..5369

5	(x:	i) SEQ	UENZBES	CHRE	BUNG	3: SI	EQ II	NO:	: 3:						•
	AAGCTT	GACC T	TGGCTGG	CA C	rtga	STCG	CA	GACA	GTG	GAC	raaco	CCG 2	AGCA	ATG	57
														Met 1	
10														_	
			TGT AAA Cys Lys												105
	ASP AIG	g GIY	5	GLY	116	361	10	VGI	neu	Der	VIG	15	vai	1110	
15	CAC ATA	A ATA	CCG ATT	ACA	ттт	GAA	ATA	TCG	ATG	GTA	TGT	GGC	ATA	TTG	153
	His Ile	e Ile	Pro Ile	Thr	Phe	Glu	Ile	Ser	Met	Val	_	Gly	Ile	Leu	
		20				25					30				
			TTT GGT												201
20	Thr Ty:		Phe Gly	Ala	Ser	Phe	Ala	Ala	Ile	Thr 45	Phe	Ser	Thr	Met	
		-													240
	= :		TCC ATC Ser Ile												249
25	50			55					60		-			65	
	TTT AG	G CGT	GAT GCG	AAC	AAG	GCT	GAC	AAT	AAG	GCC	GCT	AGT	GTG	GCA	297
			Asp Ala					Asn					Val		
			70					75					80	-	
30			CTA ATA												345
	Leu Ası	o Ser	Leu Ile 85	Asn	Phe	Glu	Ala 90	Val	Lys	Tyr	Phe	Asn 95	Asn	Glu	
												<b>ma</b> 6	222	a	202
35			GCG GAC Ala Asp												393
	-11	100	<b>-</b>			105					110	-	_	_	
	TCC CA	G ATA	AAG GTC	TCG	CAA	TCG	CTG	GCG	TTT	TTG	AAC	ACC	GGC	CAG	441
			Lys Val	Ser		Ser	Leu	Ala	Phe		Asn	Thr	Gly	Gln	
40	11!	5			120					125					
		-	TTT ACC												489
	Asn Let	n lie	Phe Thr	135	AIA	Leu	Thr	Ala	140	Met	Tyr	Met	Ala	145	
45	3.3 T. O.C.	<b>.</b>	1mc	000	mcm.	CMM	202	cmc	ccc	CAM	CMM	CITIC	mm x	አመሙ	537
	-		ATG CAG Met Gln												337
	•		150	_				155					160		
	AAT CA	A CTG	GTA TTC	CAG	CTC	TCC	GTG	CCA	CTA	AAC	TTC	CTT	GGT	AGC	585
50		_			_	_		_	_	_		-	~		

Asn Gln Leu Val Phe Gln Leu Ser Val Pro Leu Asn Phe Leu Gly Ser

170

175

165

55

				Asp										Ser		TTT Phe	633
5			Gln													CAG Gln	681
10		Leu							GAT Asp								729
15									ATA Ile								777
20									ATA Ile 250								825
05									TTT Phe								873
25									GAT Asp								921
30									GTG Val								969
35									AAA Lys								1017
40									GAA Glu 330								1065
	CTC Leu	CAG Gln	AAC Asn 340	CTA Leu	CCA Pro	AAG Lys	GGC Gly	GCT Ala 345	TCC Ser	ACC Thr	GTT Val	GTA Val	GGG Gly 350	GAG Glu	CGC Arg	GGT Gly	1113
45	Leu	ATG Met 355	ATC Ile	AGC Ser	GGA Gly	GGT Gly	GAG Glu 360	AAA Lys	CAA Gln	AGG Arg	Leu	GCT Ala 365	ATT Ile	GCT Ala	CGT Arg	GTG Val	1161
50	CTT Leu 370									Phe							1209

	CTG GAT ACA CAC ACA GAG CAG GCA CTC TTG CAC ACC ATT CAG CAG AAC Leu Asp Thr His Thr Glu Gln Ala Leu Leu His Thr Ile Gln Gln Asn 390 395 400	1257
5	TTT TCT TCC AAT TCA AAG ACG AGC GTT TAC GTT GCC CAT AGA CTG CGC Phe Ser Ser Asn Ser Lys Thr Ser Val Tyr Val Ala His Arg Leu Arg 405 410 415	1305
10	ACA ATC GCT GAT GCA GAT AAG ATC ATT GTT CTT GAA CAA GGT TCT GTC Thr Ile Ala Asp Ala Asp Lys Ile Ile Val Leu Glu Gln Gly Ser Val 420 425 430	1353
15	CGC GAA GAG GGC ACA CAC AGC TCG CTG TTA GCG TCA CAA GGA TCC CTA Arg Glu Glu Gly Thr His Ser Ser Leu Leu Ala Ser Gln Gly Ser Leu 435 440 445	1401
20	TAC CGG GGT CTG TGG GAT ATT CAG GAA AAC CTA ACG CTT CCG GAA CGG Tyr Arg Gly Leu Trp Asp Ile Gln Glu Asn Leu Thr Leu Pro Glu Arg 450 460 465	1449
	CCT GAG CAG TCA ACC GGA TCT CAG CAT GCA TAGACGTCTG ACTAGAGATT Pro Glu Gln Ser Thr Gly Ser Gln His Ala 470 475	1499
25	ATATAATAAC CCTCGAGCCA AAATTATACG GCGCTAACAA GTAAAAATTT TAGTTACTTT	1559
	TCTGACTTCT CTACGCTGAC TTCTCTACCC TTCTAACATA GTTAATTGAA GTAGTGGTTA	1619
30	ATGACGACTG CATTTTATTA TTGTCCACTT TGCATTAGAA GTACTAGTGC TTAAGCGCTC	1679 1739
	TTTAGGCCGC TTTCTTCTTC TTTGTCAGGC CGCAAGGTAA AGGAAGCACC AACGGATTGC TACCGCTGCT ATTCCTGCTC TCTCAAG ATG TGT GGC ATA TTA GGC GTT GTG	1790
35	Met Cys Gly Ile Leu Gly Val Val 1 5	
40	CTA GCC GAT CAG TCG AAG GTG GTC GCC CCT GAG TTG TTT GAT GGC TCA Leu Ala Asp Gln Ser Lys Val Val Ala Pro Glu Leu Phe Asp Gly Ser 10 15 20	1838
	CTG TTC TTA CAG CAT CGC GGT CAA GAT GCT GCC GGG ATT GCT ACG TGC Leu Phe Leu Gln His Arg Gly Gln Asp Ala Ala Gly Ile Ala Thr Cys 25 30 35 40	1886
45	GGC CCC GGT GGG CGC TTG TAC CAA TGT AAG GGC AAT GGT ATG GCA CGG Gly Pro Gly Gly Arg Leu Tyr Gln Cys Lys Gly Asn Gly Met Ala Arg 45 50 55	1934

5				Leu					Ala					Asn		GAA Glu	2030
J	GCC Ala	Gln 90	Pro	TTC Phe	TAI Tyr	GTG Val	AAT Asn 95	AGT Ser	CCC	TAC Tyr	GGA Gly	ATT 11e	Cys	ATG Met	AGT Ser	CAT His	2078
10	Asn 105	Gly	' Asn	Leu	Val	Asn 110	Thr	Met	Ser	Leu	Arg 115	Arg	Tyr	Leu	Asp	GAA Glu 120	2126
15	GAC <b>A</b> sp	GTT Val	CAC His	CGT	CAT His 125	ATT Ile	AAC Asn	ACG Thr	GAC Asp	AGC Ser 130	GAT Asp	TCT	GAG Glu	CTA Leu	CTG Leu 135	CTT Leu	2174
20	AAT Asn	ATA Ile	TTT Phe	GCC Ala 140	GCG Ala	GAG Glu	CTG Leu	GAA Glu	AAG Lys 145	TAC Tyr	AAC Asn	AAA Lys	TAT Tyr	CGT Arg 150	GTG Val	AAC Asn	2222
<i>25</i>													AAA Lys 165				2270
23	GGT Gly	GGC Gly 170	TAT Tyr	GCT Ala	TGT Cys	GTT Val	GGC Gly 175	ATG Met	TTG Leu	GCG Ala	GGA Gly	TAT Tyr 180	GGA Gly	TTG Leu	TTT Phe	GGT Gly	2318
30	TTC Phe 185	CGG Arg	GAC Asp	CCC Pro	AAT Asn	GGG Gly 190	ATC Ile	AGG Arg	CCG Pro	CTA Leu	TTG Leu 195	TTT Phe	GGT Gly	GAG Glu	CGC Arg	GTC Val 200	2366
35													GAA Glu				2414
40	CTT Leu	AAG Lys	GCC Ala	CAC His 220	CGC Arg	TTC Phe	CAA Gln	AAC Asn	ATA Ile 225	CGT Arg	GAT Asp	ATT Ile	CTT Leu	CCC Pro 230	GGC Gly	CAA Gln	2462
	GCC Ala	GTC Val	ATT Ile 235	ATC Ile	CCT Pro	AAA Lys	Thr	TGC Cys 240	GGC Gly	TCC Ser	AGT Ser	CCA Pro	CCA Pro 245	GAG Glu	TTC Phe	CGG Arg	2510
<b>4</b> 5	Gln	GTA Val 250	GTG Val	CCA Pro	ATT Ile	Glu	GCC Ala 255	TAC Tyr	AAA Lys	CCG Pro	Asp	TTG Leu 260	TTT Phe	GAG Glu	TAC Tyr	GTG Val	2558
50	TAT Tyr 265	TTC Phe	GCT Ala	CGT .	Ala	GAC Asp 270	AGC Ser	GTT Val	CTG Leu	Asp	GGT . Gly 275	ATT Ile	TCC Ser	GTT Val	TAC Tyr	CAT His 280	2606

				Gly			Glu			CAG Gln	2654
5				ATT Ile					Thr	GCA Ala	2702
10				GAG Glu							2750
15				AAC Asn							2798
20				GTA Val 350							2846
25				GAC Asp							2894
25				TCC Ser							2942
30				TAC Tyr							2990
35				ATT Ile							3038
40				GAA Glu 430							3086
				TTG Leu							3134
<b>4</b> 5		Ser		GAA Glu				Asn			3182
50	Val			TAC Tyr	Leu						3230

5	AAT AAC Asn Asn 490	Ser 2		Gly G										3278
3	TAC AAT Tyr Asn 505				AGCGGC	GCC	GTTG(	CCGG	CA TO	CGGG	CCC.	A		3326
10	TATATAG	ACT C	ATCGGGA	CC TAA	ATAAG	с ст	TTAC	AGAT	CATT	PATC	rac i	AAAT	ATAGAT	3386
	ACCATTA	AAA G	CCTGACT'	TT CGA	TTACT	C CT	AGCA	CACC	CCGT	TGT	ATC (	CTG	rgc <b>tt</b> g	3446
	CTTTCTT	AAA T	GCCGTTG	GT TAG	CTTTG	G AC	TTAG	CGTC	CCGC	CCA	rtt :	TCTAC	GCATGT	3506
15	GCAGATC	TAG C	AAATTTG(	GC CTA	GACAA	G AA	GATC	CATT	CGGC	CACC	CAC A	ATCCT	rggagc	3566
	CAGCACA	CAG TO	GGACCCA		CG AGC t Ser							-		3617
20	CTA GAG	GAG A	AAT AGC	GAA CA	G CTG	GAC	CAG	TCC	АСТ	ACG	GAG	ACT	TAC	3665
	Leu Glu													3003
25	GTG GTA Val Val													3713
	var var	Cys (	30	ASII G.	u nsp	35	Deu	ASII	GIII	FILE	40	GIII	GIII	
30	TGT TGG													3761
	Cys Trp	Gln :	Ile Asp	Glu G	y Glu. 50	_	Val	Thr	Asn	Leu 55	Glu	Pro	Leu .	
	GGA TTC										-			3809
35	Gly Phe 60		Thr Lys		l Ser 5	Arg	Asp	Glu	Glu 70	Asn	Leu	Arg	Leu	
	AAC GTA			-	-					-				3857
40	Asn Val	Tyr :	Tyr Ala	Lys Se	er Pro	Leu	Asp	Ala 85	Gln	Thr	Leu	Gln	Phe 90	
	CTG GGC	GTG T	TTC CTG	CGC C	A ATG	GAA	ACC	TCA	CAA	ATA	CGT	TGG	ATC	3905
	Leu Gly	Val I	Phe Leu 95	Arg G	n Met	Glu	Thr 100	Ser	Gln	Ile	Arg	Trp 105	Ile	
45	TTC CTA	CTG (	GAC TGG	CTG CT	'A GAC	GAT	AAA	CGA	TTA	TGG	CTA	CGT	CAA	3953
	Phe Leu	Leu A												
50	CTG CGG													4001
	Leu Arg	Asn 5	Ser Trp	Ala Al	a Leu 130		Glu	Ala	Gln	Val 135	Ala	Pro	Phe	

-			GTG Val							4049
5			ACG Thr							4097
10			CGA Arg 175							4145
15			AAT Asn						_	4193
20			CTG Leu							4241
<i>2</i> 5			ATC Ile							4289
23			GAG Glu							4337
30	 		TTC Phe 255	-					GGA Gly	4385
35			TCT Ser							4433
<b>4</b> 0			ACC Thr							4481
45			CAT His							4529
<del>10</del>			CAC His							4577
50			AAG Lys 335							4625

5	GCA GCC GAC TCA CCG AAC GAC GTC GCT GAC TCC ATC GAT GGG CTT ATG Ala Ala Asp Ser Pro Asn Asp Val Ala Asp Ser Ile Asp Gly Leu Met 350 355 360	4673
3	GAT GGT ATC GTA CAA AGG AAT GTT CAT TGACGTCGAC ACAAAAATTT Asp Gly Ile Val Gln Arg Asn Val His 365 370	4720
10	TGTTACTGTT CTCTCGAGAA CTATTCTCAT CCAGTACTGA CATATTAGAA GGCGAAGTGA	4780
	ACTAGGATTT ATATAAAGTA GCCTTCAGGC AATTGCACAG GGTCTATTGA GTCGCTGCCG	4840
	TTCACGAGAG AGCCCAATAT ATCGAGGACT AATTGGTCAC TTTTGTTTTG	4900
15	CCTGTATTTG CTAATCATTT ATCCGCTTTG TCCAAGTGGT TGCGAAGATA TCGAGCCAGA	4960
	ACATTAGAAT CTGGTTTGCC GCATCCTAGA GCTGTCTCCA AGCCAGTTGA ACCGTTGCGG	5020
	GAGATTACCG CAGCCGGTTT GATCAGAGTA CTGGTGACTG CCAGCACCCA CGTTTGTGAC	5080
20	TTATAAATAT ACGCCCTGTG GAGCCATAGC CATTGGCATA AAGAGAAGAG	5140
	CACGATGCAG ACACTTCCGG TGTACCCAGC GTCACAGACT GCGTCGCCTA CGAAGCGTGA	5200
<i>25</i>	ACTTGCAGCG GCGCCCTCGG TGCCGCAGGA CGGCGCCCGG CTGCCTGCGC AGCTCACTTT	5260
	AGTGACGCCC CCAGAACCTG ATATCCAGAA GAAGTCAGTG CGATCTCAGG TCGCGCGTTT	5320
	AAGCATCTCG GAGACAGATG TAGTGAAGAG TGATATCGTG GCTAAGCTT	5369
30	(2) INFORMATION ZU SEQ ID NO: 4:	
35	<ul><li>(i) SEQUENZ CHARAKTERISTIKA:</li><li>(A) LÄNGE: 475 Aminosäuren</li><li>(B) ART: Aminosäure</li><li>(D) TOPOLOGIE: linear</li></ul>	
<b>3</b> 5	(ii) ART DES MOLEKÜLS: Protein	
	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:	
40	Met Asp Arg Gly Cys Lys Gly Ile Ser Tyr Val Leu Ser Ala Met Val 1 5 10 15	
	Phe His Ile Ile Pro Ile Thr Phe Glu Ile Ser Met Val Cys Gly Ile 20 25 30	
<b>4</b> 5	Leu Thr Tyr Gln Phe Gly Ala Ser Phe Ala Ala Ile Thr Phe Ser Thr 35 40 45	
50	Met Leu Leu Tyr Ser Ile Phe Thr Phe Arg Thr Thr Ala Trp Arg Thr 50 55 60	
	Arg Phe Arg Arg Asp Ala Asn Lys Ala Asp Asn Lys Ala Ala Ser Val 65 70 75 80	

	Ala Leu	ı Asp Sei	Leu Il 85	e Asn	Phe	Glu	Ala 90	Val	Lys	Tyr	Phe	Asn 95	Asn
5	Glu Lys	Tyr Let		p Lys	Tyr	His 105	Thr	Ser	Leu	Met	Lys 110	Tyr	Arg
	Asp Ser	Gln Ile 115	e Lys Va	1 Ser	Gln 120	Ser	Leu	Ala	Phe	Leu 125	Asn	Thr	Gly
10	Gln Asr 130	Leu Ile	Phe Th	r Thr 135	Ala	Leu	Thr	Ala	Met 140	Met	Tyr	Met	Ala
15	Cys Asr 145	Gly Val	. Met Gl 15		Ser	Leu	Thr	Val 155	Gly	Asp	Leu	Val	Leu 160
		ı Gln Leı	165				170					175	
20		. Tyr Arc	)	_		185					190		
	-	Leu Glr 195	_		200					205			
25	210			215					220				
30	225	Gly Tyr	23	0				235					240
		Pro Ala	245	_			250					255	
35		Ser Thi 260 Arg Ile	)	_		265					270		
	_	275 Leu Arg		_	280					285		_	
40	290	)		295					300				
<b>4</b> 5	305	Asp Thi	31	0				315					320
10		Asp Glu	325				330					335	
50		Gln Asr	)			345					350		
	GIÀ PE	Met Ile 355	: ser Gl	<b>ў</b> ет <b>ў</b>	360	гуз	GIII	wrd	nen	365	TIE	AIG	NT.A

23

	Va:	1 Leu 37(	ı Lei	ı Lys	Asp	Ala	Pro 375		Met	Phe	Phe	As <sub>2</sub>		ı Ala	Th	s Sei
5	A18	a Leu	ı Asp	Thr	His	Thr 390	Glu	Glm	Ala	Leu	Leu 395	His	Thi	: Ile	Glr	Glr 400
10	Asr	ı Phe	e Ser	Ser	Asn 405	Ser	Lys	Thr	Ser	Val 410	Tyr	Val	Ala	His	415	
	Arç	Thr	Ile	Ala 420	Asp	Ala	Asp	Lys	Ile 425	Ile	Val	Leu	Glu	Gln 430		Ser
15	Val	. Arg	Glu 435	Glu	Gly	Thr	His	Ser 440	Ser	Leu	Leu	Ala	Ser 445		Gly	Ser
	Leu	450	Arg	Gly	Leu	Trp	Asp 455	Ile	Gln	Glu	Asn	Leu 460	Thr	Leu	Pro	Glu
20	Arg 465		Glu	Gln	Ser	Thr 470	Gly	Ser	Gln	His	Ala 475					
	(2)	INF	ORMA	TION	ZU S	SEQ ]	D NO	D: 5	:							
25			() ()	SEQUE A) Li B) AI D) TO	NGE:	510 mino	Ami säur	inosa ce		ı						
30			) AR	r des Quenz	MOL	EKÜL	s: E	Prote		NO.	5.					
35	Met 1			Ile								Gln	Ser	Lys	Val 15	Val
	Ala	Pro	Glu	Leu 20	Phe	Asp	Gly	Ser	Leu 25	Phe	Leu	Gln	His	Arg 30	Gly	Gln
40	Asp	Ala	Ala 35	Gly	Ile	Ala	Thr	Cys 40	Gly	Pro	Gly	Gly	Arg 45	Leu	Tyr	Gln
	Cys	Lys 50	Gly	Asn	Gly	Met .	Ala 55	Arg	Asp	Val	Phe	Thr 60	Gln	Ala	Arg	Met
45	Ser 65	Gly	Leu	Val	Gly	Ser 1	Met	Gly	Ile	Ala	His 75	Leu	Arg	Tyr	Pro	Thr 80
50	Ala	Gly	Ser	Ser .	Ala . 85	Asn :	Ser	Glu	Ala	Gln :	Pro :	Phe	Tyr	Val	Asn 95	Ser
	Pro	Tyr	Gly	Ile ( 100	Cys 1	Met :	Ser :		Asn 105	Gly A	Asn 1	Leu	Val	Asn 110	Thr	Met

	Ser	Leu	Arg 115	Arg	Tyr	Leu	Asp	Glu 120	Asp	Val	His	Arg	His 125	Ile	Asn	Thr
5	Asp	Ser 130	Asp	Ser	Glu	Leu	Leu 135	Leu	Asn	Ile	Phe	Ala 140	Ala	Glu	Leu	Glu
	Lys 145	Tyr	Asn	Lys	Tyr	Arg 150	Val	Asn	Asn	Asp	Asp 155	Ile	Phe	Cys	Ala	Leu 160
10	Glu	Gly	Val	Tyr	Lys 165	Arg	Cys	Arg	Gly	Gly 170	Tyr	Ala	Cys	Val	Gly 175	Met
15	Leu	Ala	Gly	Tyr 180	Gly	Leu	Phe	Gly	Phe 185	Arg	Asp	Pro	Asn	Gly 190	Ile	Arg
	Pro	Leu	Leu 195	Phe	Gly	Glu	Arg	Val 200	Asn	Asp	Asp	Gly	Thr 205	Met	Asp	Tyr
20	Met	Leu 210	Ala	Ser	Glu	Ser	Val 215	Val	Leu	Lys	Ala	His 220	Arg	Phe	Gln	Asn
	Ile 225	Arg	Asp	Ile	Leu	Pro 230	Gly	Gln	Ala	Val	Ile 235	Ile	Pro	Lys	Thr	Cys 240
25	Gly	Ser	Ser	Pro	Pro 245	Glu	Phe	Arg	Gln	Val 250	Val	Pro	Ile	Glu	Ala 255	Tyr
	-			260			_		265				Ala	270		
30			275					280					Met 285			
35		290					295					300	Glu			
	305					310					315		Leu			320
40					325					330			Lys		335	
				340					345				Arg	350		
<b>4</b> 5			355					360					Lys 365			
50		370					375					380	Thr			
30	Ile 385	Val	Asn	Met	Ala	Lys 390	Glu	Ser	Gly	Ala	Ala 395	Lys	Val	Tyr	Phe	Ala 400

25

	Ser	Ala	a Ala	Pro	Ala 405		Arg	Phe	Asn	His 410		Tyr	Gly	'Ile	Asp 415	Leu
5	Ala	Asp	Thr	Lys 420	Gln	Leu	Val	Ala	Tyr 425		Arg	Thr	Val	Glu 430		Ile
10	Thr	Ala	435	Leu	Gly	Cys	Asp	Arg 440	Val	Ile	Tyr	Gln	Ser 445		Asp	Asp
	Leu	11e 450	Asp	Суз	Суз	Lys	Thr 455	Asp	Ile	Ile	Ser	Glu 460		Glu	Val	Gly
15	Val 465	Phe	Thr	Gly	Asn	Tyr 470	Val	Thr	Gly	Val	Glu 475	Asp	Val	Tyr	Leu	Gln 480
	Glu	Leu	Glu	Arg	Cys 485	Arg	Ala	Leu	Asn	Asn 490	Ser	Asn	Lys	Gly	Glu 495	Ala
20	Lys	Ala	Glu	Val 500	Asp	Ile	Gly	Leu	Tyr 505	Asn	Ser	Ala	Asp	Tyr 510		
	(2)	INF	ORMA	TION	ZU S	SEQ :	ID NO	D: 6	•							
25			( <i>1</i>	SEQUE A) LÃ 3) AR D) TO	NGE:	371 mino	L Ami osāum	inosá re		ı						
30				r des Quenz							_					
35	Met 1			Gly								Glu	Glu	Asn	Ser 15	Glu
	Gln	Leu	Asp	Gln 20	Ser	Thr	Thr	Glu	Thr 25	Tyr	Val	Val	Cys	Cys 30	Glu	Asn
40	Glu	Asp	Ser 35	Leu	Asn	Gln	Phe	Leu 40	Gln	Gln	Cys	Trp	Gln 45	Ile	Asp	Glu
	Gly	Glu 50	Lys	Val	Thr	Asn	Leu 55	Glu	Pro	Leu	Gly	Phe 60	Phe	Thr	Lys	Val
45	Val 65	Ser	Arg	Asp (	Glu	Glu 70	Asn	Leu	Arg	Leu	Asn 75	Val	Tyr	Tyr	Ala	Lys 80
50	Ser	Pro	Leu	Asp 2	Ala (	Gln	Thr	Leu	Gln	Phe 90	Leu	Gly	Val	Phe	Leu 95	Arg
	Gln :	Met	Glu	Thr :	Ser (	Gln	Ile .		Trp	Ile	Phe	Leu		Asp 110	Trp	Leu

	Leu	Asp	Asp 115	Lys	Arg	Leu	Trp	Leu 120	Arg	Gln	Leu	Arg	Asn 125	Ser	Trp	Ala
5	Ala	Leu 130	Glu	Glu	Ala	Gln	Val 135	Ala	Pro	Phe	Pro	Gly 140	Gly	Ala	Val	Val
10	Val 145	Val	Leu	Asn	Pro	Ser 150	His	Val	Thr	Gln	Leu 155	Glu	Arg	Asn	Thr	Met 160
10	Val	Trp	Asn	Ser	<b>Arg</b> 165	Arg	Leu	Asp	Leu	Val 170	His	Gln	Thr	Leu	Arg 175	Ala
15	Ala	Cys	Leu	Asn 180	Thr	Gly	Ser	Ala	Leu 185	Val	Thr	Leu	Asp	Pro 190	Asn	Thr
	Ala	Arg	Glu 195	Asp	Val	Met	His	Ile 200	Cys	Ala	Leu	Leu	Ala 205	Gly	Leu	Pro
20	Thr	Ser 210	Arg	Pro	Val	Ala	Met 215	Leu	Ser	Leu	Gln	Ser 220	Leu	Phe	Ile	Pro
25	His 225	Gly	Ala	Asp	Ser	11e 230	Gly	Lys	Ile	Cys	Thr 235	Ile	Ala	Pro	Glu	Phe 240
25	Pro	Val	Ala	Thr	Val 245	Phe	Asp	Asn	Asp	Phe 250	Val	Ser	Ser	Thr	Phe 255	Glu
30	Ala	Ala	Ile	Ala 260	Pro	Glu	Leu	Thr	Pro 265	Gly	Pro	Arg	Val	Pro 270	Ser	Asp
	His	Pro	Trp 275	Leu	Thr	Glu	Pro	Thr 280	Asn	Pro	Pro	Ser	Glu 285	Ala	Thr	Ala
35	Trp	His 290	Phe	Asp	Leu	Gln	Gly 295	Arg	Leu	Ala	Thr	Leu 300	Tyr	Arg	His	Leu
40	Gly 305	Asp	Ser	Asn	Lys	Ala 310	Ile	Ser	Val	Thr	Gln 315	His	Arg	Phe	His	Lys 320
40	Pro	Arg	Ser	Glu	Asp 325		Ala	Tyr	Glu	Phe 330		Leu	Pro	Ser	<b>Lys</b> 335	His
45	Pro	Thr	Ile	Arg 340	Asp	Leu	Ile	Arg	Ser 345	Ala	Ala	Ala	Asp	Ser 350	Pro	Asn
	Asp	Val	Ala 355	Asp	Ser	Ile	Asp	Gly 360	Leu	Met	Asp	Gly	Ile 365	Val	Gln	Arg
50	Asn	Val 370	His				•									
	(2)	INF	ORMA!	rion	ZU S	SEQ I	ED NO	): 7:	:							

5	(1) SEQUENZ CHARAKTERISTIKA:  (A) LÄNGE: 3616 Basenpaare  (B) ART: Nukleinsäure  (C) STRANGFORM: Einzel  (D) TOPOLOGIE: linear	
	(ii) ART DES MOLEKÜLS: DNS (genomisch)	
10	(iii) HYPOTHETISCH: NEIN	
	(iii) ANTISENSE: NEIN	
15	(ix) MERKMALE:  (A) NAME/SCHLÜSSEL: 5'UTR  (B) LAGE: 1863	
20	(ix) MERKMALE:  (A) NAME/SCHLÜSSEL: CDS  (B) LAGE: 8641316	
	(ix) MERKMALE:  (A) NAME/SCHLÜSSEL: intron  (B) LAGE: 13171477	
25	(ix) MERKMALE:  (A) NAME/SCHLÜSSEL: CDS  (B) LAGE: 14782592	
30	(ix) MERKMALE:  (A) NAME/SCHLÜSSEL: 3'UTR  (B) LAGE: 25933616	
	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:	
35	GGGCCCGGTG CCAGCTCGCC AGGTGCGGAC TCGCGCTCGG GCTGTGGGCG CTCTACCTGC	60
	TGCTGCTCGG CAGCTGCCTG ACGCGCGCGT ACGAGCTGTC GGATCTCGAA AACCTGGAAT	120
	CCGATTACTA CAGCTACGTG CTGGATGTGA ACTTCGCGCT GCTGAGCGCC ATGAGCGCGA	180
40	CCGGCCTCGC GATGGGCGCC GTGAGCGGCT CCCTCGGGAG CGCGCCGGTG CTCGCGCAGT	240
	GGCCGGCAGC GATCTGGGCC GTGCGCTTCC TGCGCGCCGC GGGCTATGTC GCGATAGTCC	300
45	TAATCCTGCC GTTCCTGTCC GTCGTCGCAT TCCTGCAGCC GCTCTGCGAG CGCGCGCTGG	360
<b>4</b> 5	CGCTGTTCCC GTTTGTGCGC GCGTGGGGCA TGGACGGCGT GTTCAACTTC CTGCTGCTCT	420
	CCGCCGTGCT CTGGACTGTA TTCCTGGCCG TTCGCCTGCT CCGCGCCGTC TACAGACTGC	480
50	TGCGCTGGCT GGTCGGTCTT TTGGTCCGCC TGGCACGCCT GCTGCTGCGA GGCGCCCGTC	540
	GGACGCCTGC GGCGGCCCCC GAGGAGCCCG TCTAGCGTGC GCGCGTTCTA GGCCCCTGAC	600
	AGCTCCTACC TGGTGCTGGC CGCCGGTAGG GCTCGCATCG TGCGGCGCAG GCCCATTGCT	660

	TTTTGGCCCC CGCTGGATCA TCGTTTCTTT TACGTGAAAA GTTTGCAGCG ATGAGCTGCA	720
	GTATAAATAG GTTTTCTAGA TGCGCCAAAT CCCAGCTGGG TTTACCGGCG TCTGTTCGGG	780
5	ATAGTTACTT GATGGATGGG TCAACTTGAG AGCTTGGGTT TAGTGTTGAC TCCTTCTCTT	840
	CATAGCACGC CGAACAAAGC GCA ATG ACT TAC AGA GAC GCA GCC ACG GCA  Met Thr Tyr Arg Asp Ala Ala Thr Ala  1 5	890
10	CTG GAG CAC CTG GCG ACG TAC GCC GAG AAG GAC GGG CTG TCC GTG GAG Leu Glu His Leu Ala Thr Tyr Ala Glu Lys Asp Gly Leu Ser Val Glu 10 15 20 25	938
15	CAG TTG ATG GAC TCC AAG ACG CGG GGC GGG TTG ACG TAC AAC GAC TTC Gln Leu Met Asp Ser Lys Thr Arg Gly Gly Leu Thr Tyr Asn Asp Phe 30 35 40	986
20	CTG GTC TTG CCG GGC AAG ATC GAC TTC CCA TCG TCG GAG GTG GTG CTG Leu Val Leu Pro Gly Lys Ile Asp Phe Pro Ser Ser Glu Val Val Leu 45 50 55	1034
25	TCG TCG CGC CTG ACC AAG AAG ATC ACC TTG AAC GCG CCG TTT GTG TCG  Ser Ser Arg Leu Thr Lys Lys Ile Thr Leu Asn Ala Pro Phe Val Ser  60 65 70	1082
	TCG CCG ATG GAC ACG GTG ACG GAG GCC GAC ATG GCG ATC CAC ATG GCG Ser Pro Met Asp Thr Val Thr Glu Ala Asp Met Ala Ile His Met Ala 75 80 85	1130
30	CTC CTG GGC GGC ATC GGG ATC ATC CAC CAC AAC TGC ACT GCG GAG GAG Leu Leu Gly Gly Ile Gly Ile His His Asn Cys Thr Ala Glu Glu 90 95 100 105	1178
35	CAG GCG GAG ATG GTG CGC CGG GTC AAG AAG TAC GAA AAC GGG TTC ATC Gln Ala Glu Met Val Arg Arg Val Lys Lys Tyr Glu Asn Gly Phe Ile 110 115 120	1226
40	AAC GCC CCC GTG GTC GTG GGG CCG GAC GCG ACG GTG GCG GAC GTG CGC Asn Ala Pro Val Val Val Gly Pro Asp Ala Thr Val Ala Asp Val Arg 125 130 135	1274
45	CGG ATG AAC AAC GAG TTT GGG TTT GCA GGA TTT CCT GTG ACA Arg Met Lys Asn Glu Phe Gly Phe Ala Gly Phe Pro Val Thr 140 145 150	1316
	GGTATGTTAG AGTGGCACGC GGGGCTGCAC GCTGGGATGA TGATCATAAA TCAATAACTT	1376
50	TCGTTCTACT GACTGCGATC AAACGATCGT GTAGACACCT TTTACTCTGA CCGCAGACGT	1436
50	GCAGCGCCTT TTTGGCAGGA ACATGTACTA ACACATCAGC A GAT GAT GGC AAG Asp Asp Gly Lys 1	1489

	CCG Pro	Thi	GGC Gly	AAC Lys	G CTG	CAG Gln 10	Gly	ATC	C ATO	C ACC	G TCC Ser 15	Arc	F GAG	C ATO	C CAC	FTTT Phe 20	1537
5	GTC Val	GA0	GAC Asp	GAG Glu	ACC Thr 25	Leu	CTT Leu	GTG Val	TCT Ser	GAG Glu	Ile	ATC Met	G ACC	C AAC	GAC Asp 35	GTC Val	1585
10	ATC Ile	ACT Thr	GGG Gly	AAG Lys 40	Gln	GGC Gly	ATC Ile	AAC Asn	CTC Leu 45	Glu	GAG Glu	GCG Ala	AAC Asn	CAG Gln 50	Ile	CTG Leu	1633
15	AAG Lys	AAC Asn	ACC Thr 55	Lys	AAG Lys	GGC Gly	AAG Lys	CTG Leu 60	CCA Pro	ATT Ile	GTG Val	GAC Asp	GAG Glu 65	Ala	GGC Gly	TGC Cys	1681
20	CTG Leu	GTG Val 70	TCC Ser	ATG Met	CTT Leu	TCG Ser	AGA Arg 75	ACT Thr	GAC Asp	TTG Leu	ATG Met	AAG Lys 80	Asn	CAG Gln	TCC Ser	TAC Tyr	1729
25	CCA Pro 85	TTG Leu	GCC Ala	TCC Ser	AAG Lys	TCT Ser 90	GCC Ala	GAC Asp	ACC Thr	AAG Lys	CAG Gln 95	CTG Leu	CTC Leu	TGT Cys	GGT Gly	GCT Ala 100	1777
20	GCG Ala	ATC Ile	GGC Gly	ACC Thr	ATC Ile 105	GAC Asp	GCG Ala	GAC Asp	AGG Arg	CAG Gln 110	AGA Arg	CTG Leu	GCG Ala	ATG Met	CTG Leu 115	GTC Val	1825
30	GAG Glu	GCC Ala	GGT Gly	CTG Leu 120	GAC Asp	GTT Val	GTT Val	GTG Val	CTA Leu 125	GAC Asp	TCC Ser	TCG Ser	CAG Gln	GGT Gly 130	AAC Asn	TCG Ser	1873
35	GTC Val	TTC Phe	CAG Gln 135	ATC Ile	AAC Asn	ATG Met	ATC Ile	AAG Lys 140	TGG Trp	ATC Ile	AAG Lys	GAG Glu	ACC Thr 145	TTC Phe	CCA Pro	GAC Asp	1921
40	CTG Leu	CAG Gln 150	GTC Val	ATT Ile	GCT Ala	GGC Gly	AAC Asn 155	GTG Val	GTC Val	ACC Thr	AGA Arg	GAG Glu 160	CAG Gln	GCT Ala	GCC Ala	AGC Ser	1969
	TTG Leu 165	ATC Ile	CAC His	GCC Ala	Gly	GCA Ala 170	GAC Asp	GGG Gly	TTG Leu	CGT Arg	ATC Ile 175	GGT Gly	ATG Met	GGC Gly	TCT Ser	GGC Gly 180	2017
<b>4</b> 5	TCC Ser	ATC Ile	TGT Cys	Ile	ACT Thr 185	CAG Gln	GAG Glu	GTG Val	ATG Met	GCC Ala 190	TGT Cys	GGT Gly	AGA Arg	CCA Pro	CAG Gln 195	GGT Gly	2065
50	ACC (	GCT Ala	Val	TAC Tyr 200	AAC Asn	GTC . Val	ACG (	Gln	TTC Phe 205	GCC Ala	AAC Asn	CAG Gln	Phe	GGT Gly 210	GTG Val	CCA Pro	2113

								CAG Gln 220									2161
5								GTC Val									2209
10 .								TAC Tyr									2257
15								TCC Ser									2305
20								TCC Ser									2353
								ACT Thr 300									2401
25								CTG Leu									2449
30								CTA Leu									2497
35								TTC Phe									2545
40								TCC Ser								rgagtgc	2597
	CAC	TAGG	ccc i	ACAC'	ATAT	GA A	GTGG	ATCC	G GG(	CGCG	ATGG	CAC	CCAT	ACT '	TTTA'	TATTAT	2657
																PATGCA	2717
45																CAGCTC	2777
																CTGGCA	2837
50																CTCTTC	2957
50																GAAGTT	3017

	AG	ACAG	CGCC	TCG	TTCA	.GAC	CTTC	AGAC	CC G	CGTG	ACAG	C GC	TCCA	CGAC	GCA	GCACGC	С	3077
	AG	AATT	CATT	GTT	TTTA	GGT	ACTG	CACC	TT A	TCGC	TCTC	т тс	тстс	AACA	CGC	TATACA	r	3137
5	TC	GGGA	AACC	TTG	GCAA	TCG	CCAA	TATT	TT A	CTGC	GTAG	T GC	ACGC	CGTT	TTG	CATCATO	2	3197
	GT	CCAG.	AATA	GAC	CGTT	TTT	TCTT	CGAT	TT C	TTGG	AGCC.	A GG	TATA	ACAG	TTA	CAACCTO	3	3257
10	CT	CAGT	GTTT	TTG	GACT'	TCA .	ATGT.	AGCA	CC T	AAGT	CCTC	C CT	TATA	ACAA	AAG	TCTCTTC	2	3317
	CT	CCAA'	rtct	TCT	rcag'	TAC	AAAT	GTTT	AA T	ATCG.	AAAC	C AA	CATT	TCAG	TCA	CTTTCTC	2	3377
	GC	CAAC	TAAA	GGC	AAAG	ACC	AGGT	GAAT	AC G	rcca:	<b>r</b> gaa <i>i</i>	A TT	CGGT	AACC	AAT	ACGGATG	;	3437
15	CTO	GTGA(	CATG	TTA	ATT	GTC :	raat(	GTTC2	AT AZ	ACGT	PATCO	C GAG	STAT'	TTTA	GGA	CGCGGC	:	3497
	CTT	rgtt	CTTG	TAAC	TGT	CCA I	AGTA	STTGO	G T	GCGC!	rgaac	C AAC	GTA	AGTA	AAC	<b>PAGGAAA</b>	L	3557
	GCC	CCAG	ATTC	TTGG	TATT	CT 1	rgtac	CATTO	T GI	ragco	CCTG	A TCT	TGG	3CTT	CGT	GGCCC		3616
20	(2)	INF	ORMA	MOITA	ZU	SEQ	ID N	10: 8	3:									
25			) ( ) AR	SEQU A) L B) A D) T	ÄNGE RT: OPOL S MO	E: 15 Amin OGIE	61 Am nosāu E: li ULS:	inos re near Prot	äure ein	en								
30	Wa b			QUEN														
30	1	Int	туг	Arg	Asp 5		Ala	Thr	Ala	Leu 10		His	Leu	Ala	Thr 15	Tyr		
35	Ala	Glu	Lys	Asp 20	Gly	Leu	Ser	Val	Glu 25	Gln	Leu	Met	Asp	Ser		Thr		
35	Arg	Gly	Gly 35	Leu	Thr	Tyr	Asn	Asp 40	Phe	Leu	Val	Leu	Pro 45	Gly	Lys	Ile		
40	Asp	Phe 50	Pro	Ser	Ser	Glu	Val 55	Val	Leu	Ser	Ser	Arg 60	Leu	Thr	Lys	Lys		
	Ile 65	Thr	Leu	Asn	Ala	Pro 70	Phe	Val	Ser	Ser	Pro 75	Met	Asp	Thr	Val	Thr 80		
45	Glu	Ala	Asp	Met	Ala 85	Ile	His	Met	Ala	Leu 90	Leu	Gly	Gly	Ile	Gly 95	Ile		
	Ile	His	His	Asn 100	Cys	Thr	Ala	Glu	Glu 105	Gln	Ala	Glu	Met	Val 110	Arg	Arg		
50	Val	Lys	Lys 115	Tyr	Glu	Asn	Gly	Phe 120	Ile	Asn	Ala	Pro	Val 125	Val	Val	Gly		

	Pro	Asp 130	Ala	Thr	Val	Ala	<b>Asp</b> 135	Val	Arg	Arg	Met	Lys 140	Asn	Glu	Phe	Gly
5	Phe 145	Ala	Gly	Phe	Pro	Val 150	Thr									
	(2)	INFO	RMA	NOI	ZU S	SEQ I	D N	): 9:	:							
10		(	( <i>I</i>	A) LÄ B) AF	NGE:	: 371 Amino				ı						
15		(ii)	ART	DES	MOI	EKÜI	Ls: 1	Prote	ein							
		(xi)	SEÇ	QUENZ	ZBESC	CHRE	BUNG	3: SI	EQ II	NO:	9:					
20	Asp 1	Asp	Gly	Lys	Pro 5	Thr	Gly	Lys	Leu	Gln 10	Gly	Ile	Ile	Thr	Ser 15	Arg
	Asp	Ile	Gln	Phe 20	Val	Glu	Asp	Glu	Thr 25	Leu	Leu	Val	Ser	Glu 30	Ile	Met
25	Thr	Lys	Asp 35	Val	Ile	Thr	Gly	Lys 40	Gln	Gly	Ile	Asn	Leu 45	Glu	Glu	Ala
	Asn	Gln 50	Ile	Leu	Lys	Asn	Thr 55	Lys	Lys	Gly	Lys	Leu 60	Pro	Ile	Val	Àsp
30	Glu 65	Ala	Gly	Cys	Leu	Va1 70	Ser	Met	Leu	Ser	Arg 75	Thr	Asp	Leu	Met	Lys
	Asn	Gln	Ser	Tyr	Pro 85	Leu	Ala	Ser	Lys	Ser 90	Ala	Asp	Thr	Lys	Gln 95	Leu
35	Leu	Суз	Gly	Ala 100	Ala	Ile	Gly	Thr	11e 105	Asp	Ala	Asp	Arg	Gln 110	Arg	Leu
40	Ala	Met	Leu 115	Val	Glu	Ala	Gly	Leu 120	Asp	Val	Val	Val	Leu 125	Asp	Ser	Ser
	Gln	Gly 130	Asn	Ser	Val	Phe	Gln 135	Ile	Asn	Met	Ile	Lys 140	Trp	Ile	Lys	Glu
45	Thr 145		Pro	Asp	Leu	Gln 150	Val	Ile	Ala	Gly	Asn 155	Val	Val	Thr	Arg	G1u 160
	Gln	Ala	Ala	Ser	Leu 165	Ile	His	Ala	Gly	Ala 170	Asp	Gly	Leu	Arg	Ile 175	Gly
50	Met	Gly	Ser	Gly 180	Ser	Ile	Cys	Ile	Thr 185	Gln	Glu	Val	Met	Ala 190	Cys	Gly

	Arg	Pro	Gln 195	Gly	Thr	Ala	Val	Туг 200	Asn	Va1	Thr	Gln	Phe 205		Asn	Glr
5	Phe	Gly 210	Val	Pro	Суѕ	Ile	Ala 215	Asp	Gly	Gly	Val	Gln 220	Asn	Ile	Gly	His
	Ile 225	Thr	Lys	Ala	Ile	Ala 230	Leu	Gly	Ala	Ser	Thr 235	Val	Met	Met	Gly	Gly 240
10	Met	Leu	Ala	Gly	Thr 245	Thr	Glu	Ser	Pro	Gly 250	Glu	Tyr	Phe	Phe	<b>Ar</b> g 255	Asp
15	Gly	Lys	Arg	Leu 260	Lys	Thr	Tyr	Arg	Gly 265	Met	Gly	Ser	Ile	Asp 270	Ala	Met
	Gln	Lys	Thr 275	Asp	Val	Lys	Gly	Asn 280	Ala	Ala	Thr	Ser	Arg 285	туг	Phe	Ser
20	Glu	Ser 290	Asp	Lys	Val	Leu	Val 295	Ala	Gln	Gly	Val	Thr 300	Gly	Ser	Val	Ile
	Asp 305	Lys	Gly	Ser	Ile	Lys 310	Lys	Tyr	Ile	Pro	Tyr 315	Leu	Tyr	Asn	Gly	Leu 320
25	Gln	His	Ser	Cys	Gln 325	Asp	Ile	Gly	Val	<b>Ar</b> g 330	Ser	Leu	Va1	Glu	Phe 335	Arg
30	Glu	Lys	Val	Asp 340	Ser	Gly	Ser	Val	Arg 345	Phe	Glu	Phe	Arg	Thr 350	Pro	Ser
30	Ala	Gln	Leu 355	Glu	Gly	Gly		His 360	Asn	Leu	His		Tyr 365	Glu	Lys	Arg
35	Leu	Phe 370	Asp													
	(2) INFORMATION ZU SEQ ID NO: 10:															
40		(i)	(A (B (C	) LÄ: ) AR' ) ST:	CHA NGE: I': N' RANGI POLO	269 ukle FORM	7 Ba insä : Ei:	senp ure nzel								
45		(ii)						NS (	genoi	misc	h)					
		iii)					EIN									
	(:	iii)	ANT:	ISEN	SE: 1	VEIN										
50		(ix)	(A)	NAI	E: ME/SC SE: 1			: 5′t	JTR							

(ix) MERKMALE:

	(A) NAME/SCHLÜSSEL: CDS (B) LAGE: 4562033	
5	(ix) MERKMALE: (A) NAME/SCHLÜSSEL: 3'UTR (B) LAGE: 20342697	
10	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:	
	ATCGATTTCA GGAGATTTTT GGTAGCATTA TTGAGGTCAT TAGAGGCGTT CTGTGACTTT	60
	CGACGATTTG CACGCGCAGA AGAGGGCGTT CAACCAGCCT TTCGGATATT CCGGTTCGAG	GGATATT CCGGTTCGAG  TAATAAG AGGAGCAGGT  AGGCGGT TAGAGGGTGC  TGTGAAC AAGCACACGC  TGCGCTC GAGAAGCGCT  AAAACCA GCTGCCTACA  GCT GTT GAA CAA Ala Val Glu Gln  5  TTC GGG TCC CAG Phe Gly Ser Gln  20  AAT GTG TAC GCG ASN Val Tyr Ala 35  GGC TGG AAG CCA Gly Trp Lys Pro  TAC GCG GCA GAT Tyr Ala Ala Asp  70  GTT CCA ATT CTG Val Pro Ile Leu 85
15	TTATACCAGC AGGGATCAGC GCAGGCACTA GAGTGGCGGG TGCTAATAAG AGGAGCAGGT	180
	CCTGGAACTG AAGTTGCAAG AGATAAGCAT TGCGCGGAGA AGGAGGCGGT TAGAGGGTGC	240
20	AAGCGAGCAG GATGGGGTCT TCGATGAACT TCCCGTCTGG GTATGTGAAC AAGCACACGC	300
20	TGCAGGCACA CCGGTAGGGC GAGTGCAGGG TGAAAAATAT ATATGCGCTC GAGAAGCGCT	360
	GGGGATGAGT TCGTCTGCAA CGGCAGGCGG ATCTTCATCT GACAAAACCA GCTGCCTACA	420
25	TCAGTGCGAA GCTGTTCAGT GATAGAATAG GAGTA ATG GCT GCT GAT GAA CAA Met Ala Ala Val Glu Gln 1 5	473
30	GTT TCT AGC GTG TTT GAC ACC ATT TTG GTG CTG GAC TTC GGG TCC CAG Val Ser Ser Val Phe Asp Thr Ile Leu Val Leu Asp Phe Gly Ser Gln 10 15 20	521
<i>35</i>	TAC TCG CAT CTG ATC ACG CGG CGG CTG CGT GAG TTT AAT GTG TAC GCG Tyr Ser His Leu Ile Thr Arg Arg Leu Arg Glu Phe Asn Val Tyr Ala 25 30 35	569
40	GAG ATG CTT CCG TGT ACG CAG AAG ATC AGC GAG CTG GGC TGG AAG CCA Glu Met Leu Pro Cys Thr Gln Lys Ile Ser Glu Leu Gly Trp Lys Pro 40 45 50	617
40	AAG GGT GTG ATT TTG TCA GGC GGG CCG TAC TCC GTG TAC GCG GCA GAT Lys Gly Val Ile Leu Ser Gly Gly Pro Tyr Ser Val Tyr Ala Ala Asp 55 60 65 70	665
45	GCT CCG CAC GTG GAC CGG GCG GTG TTC GAG TTG GGC GTT CCA ATT CTG Ala Pro His Val Asp Arg Ala Val Phe Glu Leu Gly Val Pro Ile Leu 75 80 85	713
50	GGC ATC TGC TAC GGG CTA CAG GAG CTT GCG TGG ATA GCC GGC GCA GAG Gly Ile Cys Tyr Gly Leu Gln Glu Leu Ala Trp Ile Ala Gly Ala Glu	761

	GTG	GGG	G CGC	C GGC	C GAG	AAG	CGC	GAG	TAC	GGG	G CGC	GCC	G AC	S CTO	G CAG	C GTG		809
	Val	. Gly	Arç 105	, G17	/ Glu	Lys	Arg	Glu 110	Туг	Gly	Arg	Ala	Thr 115	Let	His	val		003
5	GAG	GAC	AGC	GCG	TGC	CCG	CTG	TTC	AAC	AAC	GTG	GAC	AGC	AGO	ACC	GTG		857 857
		120	•				125					130	)			Val		
10	TGG	ATG	TCG	CAC	GGT	GAC	AAG	CTG	CAC	GCA	CTA	CCI	GCG	GAT	TTC	CAC		905
	135	Mec	Ser	nis	GIY	140	гÀг	rea	HIS	Ala	145		) Ala	Asp	Phe	His 150	٠	
15	GTC Val	ACT	GCG	ACG	ACG	GAG	AAC	TCT	CCT	TTC	TGC	GGG	ATT	GCA	CAC	GAC		953
15	Vai	****	AIG	1111	155	GIU	ASI	ser	Pro	160	Суѕ	Gly	' Ile	Ala	His 165	Asp		
	TCG	AAG	CCA	ATC	TTC	GGG	ATC	CAG	TTC	CAC	CCT	GAG	GTG	ACG	CAC	TCC		1001
20	ser	гÃа	PIO	170	ьче	GIY	iie	GIn	Phe 175	His	Pro	Glu	Val	Thr 180	His	Ser		
	TCG	CAG	GGG	AAG	ACG	TTG	CTG	AAG	AAC	TTT	GCG	GTG	GAG	ATC	TGC	CAG		1049
	JCI	GIII	185	БУБ	IIII	Leu	Leu	190	Asn	Pne	Ala	Val	Glu 195	Ile	Суз	Gln		
25	GCC	GCG	CAG	ACC	TGG	ACG	ATG	GAA	AAC	TTC	ATT	GAC	ACC	GAG	ልሞሮ	CAG		1097
	Ala	Ala 200	Gln	Thr	Trp	Thr	Met 205	Glu	Asn	Phe	Ile	Asp 210	Thr	Glu	Ile	Gln		1037
30	CGG	ATC	CGG	ACC	CTT	GTG	GGC	ccc	ACC	GCG	GAA	GTC	ATC	GGT	GCT	GTG		1145
	Arg 215	Ile	Arg	Thr	Leu	Val 220	Gly	Pro	Thr	Ala	Glu 225	Val	Ile	Gly	Ala			
		GGC	CCT	CTTC	CAC		100									230	,	
35	Ser	Gly	Gly	Val	Asp	Ser	Thr	Val	Ala	GCG Ala	AAG Lys	CTG Leu	ATG Met	ACC	GAG Glu	GCC Ala		1193
					235					240					245			
	ATC	GGC	GAC	CGG	TTC	CAC	GCG	ATC	CTG	GTC	GAC	AAC	GGT	GTT	CTG	CGC		1241
40	116	GIY	Asp	250	Phe	His	Ala	Ile	Leu 255	Val	Asp	Asn	Gly	Va1 260	Leu	Arg		
	CTC	AAC	GAA	GCG	GCC	AAT	GTG	AAG	AAA	ATC	CTC	GGC	GAG	GGC	TTG	GGC		1289
	Leu	ASII	265	АІа	Ата	Asn		Lys 270	Lys	Ile	Leu	Gly	Glu 275	Gly	Leu	Gly		
45	ATC	AAC	TTG	ACT	GTT	GTT	GAC	GCC	TCC	GAA	GAG	TTC	ТТG	ACG	AAG	CTC		1337
	116	Asn 280	ьeu	Thr	Val		Asp . 285	Ala	Ser	Glu		Phe 290	Leu	Thr	Lys	Leu		
50	AAG	GGC	GTC	ACG	GAC	CCT	GAG :	AAG .	AAG .	AGA .	AAG	ATC	ATC	GGT	AAC	ACC	:	1385
	Lys (	чТĀ	val	Thr .	Asp	Pro ( 300	Glu 1	Lys :	Lys .		Lys 305	Ile	Ile	Gly	Asn			
																310		

						GCC Ala 320					1433
5						GGT Gly		_			1481
10						TCT Ser					1529
15						ATG Met			 	·	1577
20						GTG Val					1625
						AGA Arg 400					1673
25						GTC Val					1721
30					•	ATC Ile					1769
35						TTT Phe					1817
40		_				AGA Arg					1865
						TTC Phe 480					1913
<b>4</b> 5						GTC Val					1961
50						TAC Tyr					2009

_	CCA GCT ACC GTT GAA TGG GAA TAATCACCCT TGGGATCCGC TGACTGGCTA Pro Ala Thr Val Glu Trp Glu 520 525	2060
5	CTGTAATTCT ATGTAGTGGA TTAGTACGAT AAGTTACTTT TGTATGATAG ATGTAATCAC	2120
	ATCTGGCTAT TAAAATGACT CAGCCGAGGT AAATCTAACG TCCCTTCACA AGGGTGTTCC	2180
10	TGTGTGGACT TCCGCCTGAA TTTTTATAGA TATATAGATA CTCTACTCAT GAACAACCTG	2240
	CAACCGAATA AGCATTAGTG CCAGGAGAAG AGAACCGTGG AAATGGGGCA AGTAGAAAAA	2300
	ATCATATTCC TTAAGAATAA GACAGTACCA GAGGACCATT ACGAGACGAT TTTTGAATCG	2360
15	AATGGCTTCC AGACTCACTT TGTACCCATA ATAACCCATG AACACCTGCC AGATGAGGTT	2420
	CGCGGTCGAC TATCCGACGC GAATTACATG AAAAGGTTGA ATTGTTTGGT GGTAACCTCT	2480
	CAGAGGACTG TGGAGTGTCT CTATGAGGAC GTTCTGCCCT CTCTTCCAGC TGAAGCACGC	2540
20	AAATCTCTTC TCAATACGCC AGTATTCGTG GTTGGGCGTG CCACTCAGGA ATTTATGGAG	2600
	AGATGCGGCT TTACGGACGT GAGAGGGGGA TCTGAGACTG GTAATGGCGT TTTGCTAGCG	2660
<b>0</b> E	GAGTTAATGT TAAATATGAT CCAGAAGGGC GATGGGG	2697
25	(2) INFORMATION ZU SEQ ID NO: 11:	
30	<ul><li>(i) SEQUENZ CHARAKTERISTIKA:</li><li>(A) LÄNGE: 525 Aminosäuren</li><li>(B) ART: Aminosäure</li><li>(D) TOPOLOGIE: linear</li></ul>	
	(ii) ART DES MOLEKÜLS: Protein	
35	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:	
	Met Ala Ala Val Glu Gln Val Ser Ser Val Phe Asp Thr Ile Leu Val 1 5 10 15	
40	Leu Asp Phe Gly Ser Gln Tyr Ser His Leu Ile Thr Arg Arg Leu Arg 20 25 30	
	Glu Phe Asn Val Tyr Ala Glu Met Leu Pro Cys Thr Gln Lys Ile Ser 35 40 45	
45	Glu Leu Gly Trp Lys Pro Lys Gly Val Ile Leu Ser Gly Gly Pro Tyr 50 55 60	
	Ser Val Tyr Ala Ala Asp Ala Pro His Val Asp Arg Ala Val Phe Glu 65 70 75 80	
50	Leu Gly Val Pro Ile Leu Gly Ile Cys Tyr Gly Leu Gln Glu Leu Ala 85 90 95	

38

	Trp	Ile	Ala	Gly 100	Ala	Glu	Val	Gly	Arg 105	Gly	Glu	Lys	Arg	Glu 110	Tyr	Gly
5	Arg	Ala	Thr 115	Leu	His	Val	Glu	Asp 120	Ser	Ala	Cys	Pro	Leu 125	Phe	Asn	Asn
	Val	Asp 130	Ser	Ser	Thr	Val	Trp 135	Met	Ser	His	Gly	Asp 140	Lys	Leu	His	Ala
10	Leu 145	Pro	Ala	Asp	Phe	His 150	Val	Thr	Ala	Thr	Thr 155	Glu	Asn	Ser	Pro	Phe 160
15	Cys	Gly	Ile	Ala	His 165	Asp	Ser	Lys	Pro	Ile 170	Phe	Gly	Ile	Gln	Phe 175	His
	Pro	Glu	Val	Thr 180	His	Ser	Ser	Gln	Gly 185	Lys	Thr	Leu	Leu	Lys 190	Asn	Phe
20	Ala	Val	Glu 195	Ile	Cys	Gln	Ala	Ala 200	Gln	Thr	Trp	Thr	Met 205	Glu	Asn	Phe
	Ile	Asp 210	Thr	Glu	Ile	Gln	Arg 215	Ile	Arg	Thr	Leu	Val 220	Gly	Pro	Thr	Ala
<b>2</b> 5	Glu 225	Val	Ile	Gly	Ala	Val 230	Ser	Gly	Gly	Val	Asp 235	Ser	Thr	Val	Ala	Ala 240
00	Lys	Leu	Met	Thr	Glu 245	Ala	Ile	Gly	Asp	Arg 250	Phe	His	Ala	Ile	Leu 255	Val
30	Asp	Asn	Gly	Val 260	Leu	Arg	Leu	Asn	Glu 265	Ala	Ala	Asn	Val	Lys 270	Lys	Ile
35	Leu	Gly	Glu 275	Gly	Leu	Gly	Ile	Asn 280	Leu	Thr	Val	Val	Asp 285	Ala	Ser	Glu
	Glu	Phe 290	Leu	Thr	Lys	Leu	Lys 295	Gly	Val	Thr	Asp	Pro 300	Glu	Lys	Lys	Arg
40	Lys 305	Ile	Ile	Gly	Asn	Thr 310	Phe	Ile	His	Val	Phe 315	Glu	Arg	Glu	Ala	Ala 320
	Arg	Ile	Gln	Pro	Lys 325	Asn	Gly	Glu		Ile 330	Glu	Phe	Leu	Leu	Gln 335	Gly
45	Thr	Leu	Tyr	Pro 340	Asp	Val	Ile	Glu	Ser 345	Ile	Ser	Phe	Lys	Gly 350	Pro	Ser
50	Gln		Ile 355	Lys	Thr	His	His	Asn 360	Val	Gly	Gly	Leu	Leu 365	Asp	Asn	Met
	Lys	Leu 370	Lys	Leu	Ile		Pro 375	Leu	Arg	Glu	Leu	Phe 380	Lys	Asp	Glu	Val

	Arg His Leu Gly Glu Leu Leu Gly Ile Ser His Glu Leu Val Trp Arg 385 390 395 400
. 5	His Pro Phe Pro Gly Pro Gly Ile Ala Ile Arg Val Leu Gly Glu Val 405 410 415
	Thr Lys Glu Gln Val Glu Ile Ala Arg Lys Ala Asp His Ile Tyr Ile 420 425 430
10	Glu Glu Ile Arg Lys Ala Gly Leu Tyr Asn Lys Ile Ser Gln Ala Phe 435 440 445
15	Ala Cys Leu Leu Pro Val Lys Ser Val Gly Val Met Gly Asp Gln Arg 450 455 460
	Thr Tyr Asp Gln Val Ile Ala Leu Arg Ala Ile Glu Thr Thr Asp Phe 465 470 475 480
20	Met Thr Ala Asp Trp Tyr Pro Phe Glu His Glu Phe Leu Lys His Val 485 490 495
	Ala Ser Arg Ile Val Asn Glu Val Glu Gly Val Ala Arg Val Thr Tyr 500 505 510
25	Asp Ile Thr Ser Lys Pro Pro Ala Thr Val Glu Trp Glu 515 520 525
30	(2) INFORMATION ZU SEQ ID NO: 12:  (i) SEQUENZ CHARAKTERISTIKA:  (A) LÂNGE: 1634 Basenpaare  (B) ART: Nukleinsäure  (C) STRANGFORM: Doppel  (D) TOPOLOGIE: linear
35	(ii) ART DES MOLEKÜLS: cDNS zu mRNS
40	(iii) HYPOTHETISCH: NEIN  (iii) ANTISENSE: NEIN  (ix) MERKMALE:  (A) NAME/SCHLÜSSEL: 5'UTR  (B) LAGE: 1519
45	(ix) MERKMALE: (A) NAME/SCHLÜSSEL: CDS (B) LAGE: 5201482
50	(ix) MERKMALE: (A) NAME/SCHLÜSSEL: 3'UTR (B) LAGE: 14831634
	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:
<i>55</i>	

	CCTCGA	ACAT	CTAT	CTTC	TG A	GCTC	GATA	G TC	TACG	AAAT	CGG	CACA	CTA	GCCT	AATTGC		60
	CGAGAT	'GAAG	AGCT	'CCAG	GG A	ACCG	TTAA	A GA	TCTG	ATGT	TCC	ATCT	TCA	ATCA	GGACAA		120
5	ATGTTA	CGGG	ATGT	ссст	GA C	GCCA	CAGA	A GG	TAGC	CTGG	TGG	TCCA	GAC	AGAA	AAAGAG		180
	CCTACA	CCAA	AGAA	.GAAA	CA T	AACA	AGAA	A AA	GCCT	CCGC	ATC	GTTT	TGG	TAAA	TCATAA		240
	TAGGCA	CGAT	GCGC	ATAT	AC C	CTGA	CCAT	C AT	AGCG	GTTC	ccc	CCGC	TAA	CTGC	TCCGAG		300
10	CGGGTA	ACCC	CATG	TCAC	AA A	GTGA	СТСТ	G TC	тстт	CGTG	GTA	GGTG	ATG	TCAA	ATTTTC		360
	ACGACT	TCCC	ACCC	CGAT	GA G	CATC	CGTA	T TC	СТТТ	TCAT	CTA	AATT	CTA	ATAG	ATGGCT		420
15	TATGGA	TTCT	TATT	GGCG.	AC T	TACA	AGCC	T AT	GTAG	TTGG	СТТ	ссст	CAA	GTGT	TCGTAG		480
	TCTACC	ACCT	CACA	cccg	GT C	TAAC.	AGCT	T AC	GAGA	ATA .	ATG	GCT	ACT	AAT ·	GCA		534
										1	Met 1	Ala '	Thr .	Asn .	Ala 5		
20	ATC AA	G CTT	CTT	GCG	CCA	GAT	ATC	CAC	AGG	GGT	CTG	GCA	GAG	CTG	GTC	!	582
	Ile Ly	s Leu	Leu	Ala 10	Pro	Asp	Ile	His	Arg 15	Gly	Leu	Ala	Glu	Leu 20	Val		
	GCT AA	A CGC	СПА		ጥጥል	CCT	ርሞር	AC A		TICC	አልሮ	CTT	አአሮ		CAM		630
25	Ala Ly		Leu	Gly				Thr								,	030
			25					30					35				
	TGT AA Cys As															(	678
30		40					45		_			50	_	_			
	GAT AT															•	726
35	Asp I1		TTE	Tie	Tnr	GIn	Val	GIĀ	ser	GIA	Asp 65	Val	Asn	Asp	Arg		
	GTG CT	G GAG	CTG	CTC	ATC	ATG	ATC	AAC	GCT	AGC	AAG	ACG	GCG	TCT	GCG	•	774
	Val Le	u Glu	Leu	Leu	Ile 75	Met	Ile	Asn	Ala	Ser 80	Lys	Thr	Ala	Ser	Ala 85		
40	CGG CG.	ል ልጥጥ	ACG	CCT		ልጥጥ	CCA	ልልሮ	ጥጥር		ጥልር	GCG	ccc	CAG		•	822
	Arg Ar			Ala		_ =		_	Phe	_		_				,	942
				90					95					100			
45	CGG AA															8	870
	5 -1	<i>p</i>	105		9			110			_, 0		115				
	ATG CT	G ACT	ACC	GCG	GGC	TGC	GAT	CAT	GTC	ATC	ACC	ATG	GAC	TTA	CAC	9	918
50	Met Le	u Thr 120		Ala	Gly	Суѕ	Asp 125	His	Val	Ile	Thr	Met 130	Asp	Leu	His		

		TCG Ser 135	Gln											966
5		A GAG L Glu												1014
10		GCC Ala												1062
15		CTA Leu												1110
20		GCA Ala												1158
		GAT Asp 215												1206
25		CTG Leu												1254
30		ATA Ile											GAG Glu	1302
35		ATC Ile												1350
		TTC Phe												1398
40		TCG Ser 295												1446
<b>4</b> 5		ATC Ile							TGAT	' <b>TT</b> TG	CT I	CTCG	ATGCT	1499
50													CTATT ATATA	1559 1619
- <del>-</del>	GTT	GTTTC	AG A	GACC	2									1634

	(2)	INFO	OR <b>MA</b> T	NOI	ZU S	SEQ 1	ED NO	): 13	3:						-	
5		•	( <i>I</i>	A) LÄ B) AF	NGE:	320 Amino				n						
10		•						Prote			4.0					
		, ,	_	-				3: SE								
	Met 1	Ala	Thr	Asn	Ala 5	Ile	Lys	Leu	Leu	Ala 10	Pro	Asp	Ile	His	Arg 15	Gly
15	Leu	Ala	Glu	Leu 20	Val	Ala	Lys	Arg	Leu 25	Gly	Leu	Arg	Leu	Thr 30	qaA	Cys
20	Lys	Leu	Lys 35	Arg	Asp	Cys	Asn	Gly 40	Glu	Ala	Thr	Phe	Ser 45	Ile	Gly	Glu
_	Ser	<b>Val</b> 50	Arg	Asp	Gln	Asp	Ile 55	Tyr	Ile	Ile	Thr	Gln 60	Val	Gly	Ser	Gly
25	Asp 65	Val	Asn	Asp	Arg	Val 70	Leu	Glu	Leu	Leu	Ile 75	Met	Ile	Asn	Ala	Ser 80
	Lys	Thr	Ala	Ser	Ala 85	Arg	Arg	Ile	Thr	Ala 90	Val	Ile	Pro	Asn	Phe 95	Pro
30	Tyr	Ala	Arg	Gln 100	Asp	Arg	Lys	Asp	Lys 105	Ser	Arg	Ala	Pro	Ile 110	Thr	Ala
35	Lys	Leu	Met 115	Ala	Asp	Met	Leu	Thr 120	Thr	Ala	Gly	Cys	Asp 125	His	Val	Ile
33	Thr	Met 130	Asp	Leu	His	Ala	Ser 135	Gln	Ile	Gln	Gly	Phe 140	Phe	Asp	Val	Pro
40	Val 145	Asp	Asn	Leu	Tyr	Ala 150	Glu	Pro	Ser	Val	Val 155	Lys	Tyr	Ile	Lys	Glu 160
	His	Ile	Pro	His	Asp 165	Asp	Ala	Ile	Ile	Ile 170	Ser	Pro	Asp	Ala	Gly 175	Gly
45	Ala	Lys	Arg	Ala 180	Ser	Leu	Leu	Ser	Asp 185	Arg	Leu	Asn	Leu	Asn 190	Phe	Ala
	Leu	Ile	His 195	Lys	Glu	Arg	Ala	Lys 200	Ala	Asn	Glu	Val	Ser 205	Arg	Met	Val
50	Leu	Val 210	Gly	Asp	Val	Thr	Asp 215	Lys	Val	Cys	Ile	Ile 220	Val	Asp	Asp	Met

	Ala 225	Asp	Thr	Cys	Gly	Thr 230	Leu	Ala	Lys	Ala	Ala 235	Glu	Val	Leu	Leu	Glu 240
5	His	Asn	Ala	Arg	Ser 245	Val	Ile	Ala	Ile	Val 250	Thr	His	Gly	Ile	Leu 255	Ser
10	Gly	Lys	Ala	Ile 260	Glu	Asn	Ile	Asn	Asn 265	Ser	Lys	Leu	Asp	Arg 270	<b>Val</b>	Val
	Cys	Thr	Asn 275	Thr	Val	Pro	Phe	Glu 280	Glu	Lys	Met	Lys	Leu 285	Cys	Pro	Lys
15	Leu	Asp 290	Val	Ile	Asp	Ile	Ser 295	Ala	Val	Leu	Ala	Glu 300	Ser	Ile	Arg	Arg
20	Leu 305	His	Asn	Gly	Glu	Ser 310	Ile	Ser	Tyr	Leu	Phe 315	Lys	Asn	Asn	Pro	Leu 320

## Patentansprüche

25

30

45

- Protein mit der in SEQ ID NO:2 dargestellten Polypeptidsequenz oder einer aus SEQ ID NO:2 durch Substitution, Insertion oder Deletion von bis zu 15% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer Phosphoribosylpyrophosphat-Synthetase.
- 2. Protein nach Anspruch 1, das keine feed-back Hemmung durch Folgeprodukte von Stoffwechselwegen, die von Produkten des Enzyms ausgehen, mehr aufweist.
- 35 3. Protein nach Anspruch 1, das nicht mehr durch Zwischenprodukte der Purinbiosynthese, insbesondere durch Purinbasen, Purinnukleoside, Purinnukleotid-5'-Monophosphate oder Purinnukleotid-5'-Diphosphate oder Purinnukleotid-5'-Triphosphate gehemmt werden.
- 4. Protein nach Anspruch 1, bei dem eine oder mehrere der folgenden Aminosäureaustausche vorliegen: Lysin an Position 7 ausgetauscht gegen Valin, Aspartat an Position 52 ausgetauscht gegen Histidin, Leucin an Position 133 ausgetauscht gegen Isoleucin, Aspartat an Position 186 ausgetauscht gegen Histidin, Alanin an Position 193 ausgetauscht gegen Valin oder Histidin an Position 196 ausgetauscht gegen Glutamin.
  - 5. Nukleinsäuresequenz codierend für ein Protein gemäß Anspruch 1.
  - 6. Protein mit der in SEQ ID NO: 13 dargestellten Polypeptidsequenz oder einer aus SEQ ID NO: 13 durch Substitution, Insertion oder Deletion von bis zu 10% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer Phosphoribosylpyrophosphat-Synthetase.
- 50 7. Nukleinsäuresequenz codierend für ein Protein gemäß Anspruch 6.
  - Protein mit der in SEQ ID NO:5 dargestellten Polypeptidsequenz oder einer aus SEQ ID NO:5 durch Substitution, Insertion oder Deletion von bis zu 10% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer Glutamin-Phosphoribosylpyrophosphat-Amidotransferase.
  - 9. Protein nach Anspruch 8, das keine feed-back Hemmung durch Folgeprodukte von Stoffwechselwegen, die von Produkten des Enzyms ausgehen, mehr aufweist.

- 10. Protein nach Anspruch 8, das nicht mehr durch Zwischenprodukte der Purinbiosynthese, insbesondere durch Purinbasen, Purinnukleoside, Purinnukleotid-5'-Monophosphate oder Purinnukleotid-5'-Diphosphate oder Purinnukleotid-5'-Triphosphate gehemmt werden.
- 11. Protein nach Anspruch 8, bei dem eine oder mehrere der folgenden Aminosäureaustausche vorliegen: Aspartat an Position 310 ausgetauscht gegen Valin, Lysin an Position 333 ausgetauscht gegen Alanin oder Alanin an Position 417 ausgetauscht gegen Tryptophan.
  - 12. Nukleinsäuresequenz codierend für ein Protein gemäß Anspruch 8.
  - 13. Protein mit der in SEQ ID NO:8 und 9 dargestellten Polypeptidsequenz oder einer aus SEQ ID NO:8 und 9 durch Substitution, Insertion oder Deletion von bis zu 20% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer IMP-Dehydrogenase.
- 15 14. Protein nach Anspruch 13, das keine feed-back Hemmung durch Folgeprodukte von Stoffwechselwegen, die von Produkten des Enzyms ausgehen, mehr aufweist.
  - 15. Protein nach Anspruch 13, das nicht mehr durch Zwischenprodukte der Purinbiosynthese, insbesondere durch Purinbasen, Purinnukleoside, Purinnukleotid-5'-Monophosphate oder Purinnukleotid-5'-Diphosphate oder Purinnukleotid-5'-Triphosphate gehemmt werden.
  - 16. Nukleinsäuresequenz codierend für ein Protein gemäß Anspruch 13.
- 17. Protein mit der in SEQ ID NO:11 dargestellten Polypeptidsequenz oder einer aus SEQ ID NO:11 durch Substitution, Insertion oder Deletion von bis zu 10% der Aminosäuren erhältlichen Polypeptidsequenz und der enzymatischen Aktivität einer GMP-Synthetase.
  - 18. Protein nach Anspruch 17, das keine feed-back Hemmung durch Folgeprodukte von Stoffwechselwegen, die von Produkten der Enzyme ausgehen, mehr aufweist.
  - 19. Protein nach Anspruch 17, das nicht mehr durch Zwischenprodukte der Purinbiosynthese, insbesondere durch Purinbasen, Purinnukleoside, Purinnukleotid-5'-Monophosphate oder Purinnukleotid-5'-Diphosphate oder Purinnukleotid-5'-Triphosphate gehemmt werden.
- 20. Nukleinsäuresequenz codierend für ein Protein gemäß Anspruch 17.
  - 21. Verwendung einer oder mehrerer der Nukleinsäuresequenzen nach den oben genannten Ansprüchen zur gentechnischen Konstruktion von Mikroorganismen, die zur Herstellung von Riboflavin in der Lage sind.
- 22. Verfahren zur Herstellung von Riboflavin durch Kultivierung von Mikroorganismen, die in mindestens einem Gen der Purinbiosynthese genetisch verändert worden sind.
  - 23. Verfahren nach Anspruch 22, dadurch gekennzeichnet, daß es sich bei dem Mikroorganismus um ein Bakterium der Gattung Bacillus oder Corynebakterium handelt.
  - 24. Verfahren nach Anspruch 22, dadurch gekennzeichnet, daß es sich bei dem Mikroorganismus um einen eukaryontischen Mikroorganismus handelt.
- 25. Verfahren nach Anspruch 24, dadurch gekennzeichnet, daS es sich bei dem Mikroorganismus um Ashbya gossypii handelt.
  - 26. Verfahren nach Anspruch 24, dadurch gekennzeichnet, daß es sich bei dem Mikroorganismus um eine Hefe handelt.
- 27. Verfahren nach Anspruch 22, dadurch gekennzeichnet, daß es sich bei dem Mikroorganismus um eine Hefe der Gattung Candida, Saccharomyces oder Pichia handelt.
  - 28. Verfahren nach Anspruch 22, dadurch gekennzeichnet, daß die Veränderung in mindestens einer zusätzlichen

10

20

30

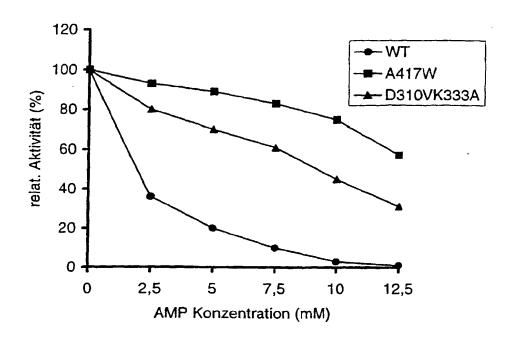




Kopie von mindestens einer der Nukleinsäuresequenzen gemäß Anspruch 5, 12, 16, 20 besteht.

29. Verfahren nach Anspruch 22, dadurch gekennzeichnet, daß durch die genetische Veränderung ein Gen codierend für ein Protein gemäß Anspruch 1, 6, 13, 17 erzeugt wird.

Abb. 1



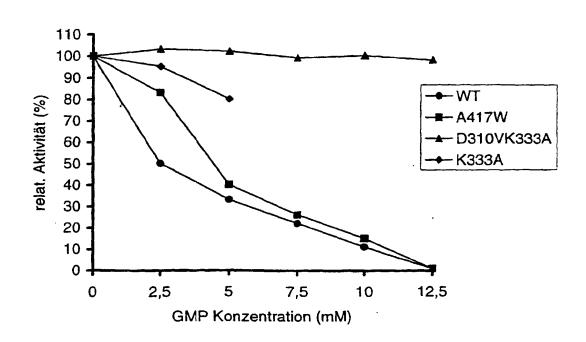
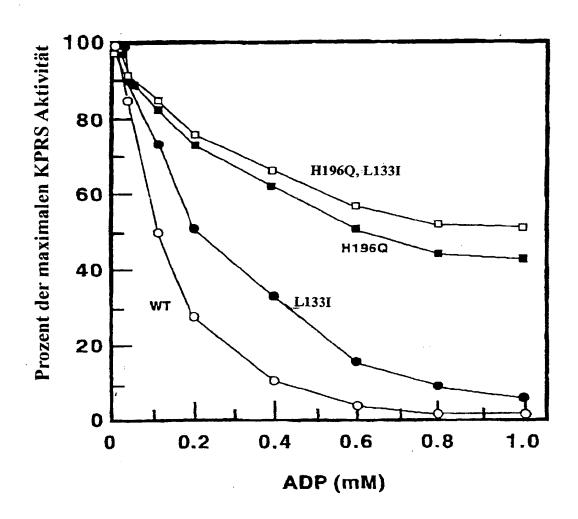


Abb. 2



Hemmung der Wildtyp und mutagenisierten KPRS durch ADP

# **Europäisches Patentamt European Patent Office** Office européen des brevets



(1) Publication number:

0 405 370 A1

(12)

## **EUROPEAN PATENT APPLICATION**

(21) Application number: 90111916.4

② Date of filing: 22.06.90

(51) Int. Cl.5; C12N 15/52, C12P 25/00, C12N 1/20, //(C12N1/20, C12R1:125,1:19)

Priority: 22.06.89 US 370378

43 Date of publication of application: 02.01.91 Bulletin 91/01

(P) Designated Contracting States: AT BE CH DE DK FR GB IT LI NL 71 Applicant: F. HOFFMANN-LA ROCHE AG Postfach 3255 CH-4002 Basel(CH)

2 Inventor: Perkins, John B. 66 Hopkins Street Reading, Massachusetts 01867(US)

Inventor: Pero, Janice G. 20 Solomon Pierce Road

Lexington, Massachusetts 02173(US)

inventor: Sloma, Alan 20 James Street

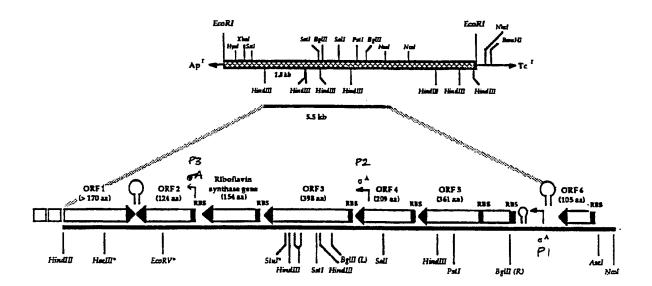
Watertown, Massachusetts 02172(US)

(72) Representative: Lederer, Franz, Dr. et al Lederer, Keller & Riederer, Patentanwälte, Lucile-Grahn-Strasse 22 22 D-8000 München 80(DE)

(54) Riboflavinoverproducing strains of bacteria.

The present invention is directed to bacteria that overproduce riboflavin. The invention concerns the nucleotide sequence of the rib operon and its open reading frames, and recombinant bacteria that contain the rib operon. Specifically, the invention relates to bacteria that have been mutated so that their production of riboflavin and/or purines is deregulated, and which have copies of the rib operon inserted and amplified in their chromosomal DNA. In a specific embodiment, the rib operon itself can be deregulated by replacing its control regions with sequences that allow constitutive or unregulated expression. The bacteria, operons and sequences of this invention are used to overproduce relatively large amounts of riboflavin by fermentation.

# FIGURE 4.



## RIBOFLAVINOVERPRODUCING STRAINS OF BACTERIA

Riboflavin (vitamin B<sub>2</sub>) is synthesized by all plants and many microorganisms but is not produced by higher animals. Because it is a precursor to coenzymes such as flavin adenine dinucleotide and flavin mononucleotide, that are required in the enzymatic oxidation of carbohydrates, riboflavin is essential to basic metabolism. In higher animals, insufficient riboflavin can cause loss of hair, inflammation of the skin, vision deterioration, and growth failure.

Riboflavin can be commercially produced either by a complete chemical synthesis, starting with ribose, or by fermentation with the fungi Eremothecium ashbyii or Ashbya gossypii (The Merck Index, Windholz et al., eds., Merck & Co., p. 1183, 1983). Mutants of Bacillus subtilis, selected by exposure to the purine analogs azaguanine and azaxanthine, have been reported to produce riboflavin in recoverable amounts (U.S. Patent No. 3,900,368, Enei et al., 1975). In general, exposure to purine or riboflavin analogs selects for deregulated mutants that exhibit increased riboflavin biosynthesis, because the mutations allow the microorganism to "compete out" the analog by increased production (Matsui et al., Agric. Biol. Chem. 46:2003, 1982). A purine-requiring mutant of Saccharomyces cerevisiae that produces riboflavin has also been reported (U.S. Patent No. 4,794,081, Kawai et al., 1988). Rabinovich et al. (Genetika 14:1696 (1978)) report that the riboflavin operon (rib operon) of B. subtilis is contained within a 7 megadalton (Md) EcoRI fragment (later referred to as a 6.3 Md fragment in Chikindas et al., Mol. Genet. Mik. Virusol. no. 2:20 (1987)). It is reported that amplification of the rib operon may have been achieved in E. coli by cloning the operon into a plasmid that conferred resistance to ampicillin and exposing bacteria containing that plasmid to increasing amounts of the antibiotic. The only evidence for rib amplification is a coincident increase in the presence of a green-fluorescing substance in the medium; the authors present a number or alternative possibilities besides an actual amplification of the operon to explain the phenomenon observed.

French Patent Application No. 2,546,907, by Stepanov et al. (published December 7, 1984), discloses a method for producing riboflavin that utilizes a mutant strain of B. subtilis which has been exposed to azaguanine and roseoflavin and that is transformed with a plasmid containing a copy of the rib operon.

Morozov et al. (Mol. Genet. Mik. Virusol. no. 7:42 (1984)) describe the mapping of the B. subtilis rib operon by assaying the ability of cloned B. subtilis rib fragments to complement E. coli riboflavin auxotrophs or to marker-rescue B. subtilis riboflavin auxotrophs. Based on the known functions of the E. coli rib genes, the following model was proposed for the B. subtilis operon: ribG (encoding a deaminase) - ribO (the control element) - ribB (a synthetase) -ribF - ribA (a GTP-cyclohydrolase) - ribT/D (a reductase and an isomerase, respectively) - ribH (a synthetase).

Morozov et al. (Mol. Genet. Mik. Virusol. no. 11:11 (1984)) describe the use of plasmids containing the B. subtilis rib operon with either wild-type (ribO ) or constitutive (ribO 335) operator regions to assay their ability to complement B. subtilis riboflavin auxotrophs. From the results, a revised model of the rib operon was proposed, with ribO now located upstream of all of the structural genes, including ribG, and with the existence of an additional operator hypothesized, possibly located just upstream of ribA.

Morozov et al. (Mol. Genet. Mik. Virusol. no. 12:14 (1985)) report that the B. subtilis rib operon contains a total of three different promoters (in addition to a fourth "promoter" that is only active in E. coli). The primary promoter of the operon was reported to be located within the ribO region, with the two secondary promoters reported between the ribB and ribF genes and within the region of the ribTD and ribH genes, respectively.

Chikindas et al. (Mol. Genet. Mik. Virusol. no. 2:20 (1987)) propose a restriction enzyme map for a 6.3 Md DNA fragment that contains the rib operon of B. subtilis. Sites are indicated for the enzymes EcoRl, Pstl, Sall, EcoRV, Pvull and HindIII.

Chikindas et al. (Mol. Genet. Mik. Virusol. no. 4:22 (1987) report that all of the structural genes of the B. subtilis rib operon are located on a 2.8 Md BgIll-HindIII fragment and that the BgIll site is located between the primary promoter of the operon and the ribosomal-binding site of its first structural gene. As described infra, Applicants show that this BgIll site is actually located within the most-5 open reading frame of the rib operon, so that the 2.8 Md fragment described does not contain all of the rib structural genes. Thus, in contrast to the report of Chikindas et al., the 1.3 Md BgIll fragment does not contain the ribosomal-binding site of the first structural gene; insertions at this site lead to a riboflavin-negative phenotype. Consequently, any attempt to use this BgIll site to engineer the rib operon in order to increase expression, for example by replacing the 5 regulatory region with a stronger promoter, would actually destroy the integrity of the first structural gene and thus the operon as well.

Chikindas et al. (Dokl. Akad. Nauk. 5 SSSR 298:997 (1988)) disclose another model of the B. subtilis rib operon, containing the primary promoter, p<sub>1</sub>, and two minor promoters, p<sub>2</sub> and p<sub>3</sub>: ribO(p<sub>1</sub>)-ribG-ribB-p<sub>2</sub>-

ribF-ribA-ribT-ribD-p<sub>3</sub>-ribH As before, it is incorrectly reported that the 1.3 Md Bglll fragment contains the entire first structural gene of the operon and that this proximal Bglll site maps within the primary regulatory region.

The present invention is directed to bacteria that overproduce riboflavin. The invention relates to the nucleotide sequence of the rib operon and its open reading frames, and recombinant bacteria that contain the rib operon. Specifically, the invention is directed to bacteria that have been mutated so that their production of riboflavin and/or purines is deregulated, and to bacteria which have copies of the rib operon inserted and amplified within their chromosomal DNA. In one embodiment, the rib operon itself can be deregulated by replacing its control regions with sequences that allow constitutive or unregulated expression. The bacteria, operons and sequences of this invention can be used to produce large amounts of riboflavin by fermentation.

The present invention is illustrated by way of specific examples detailed below, for example, a mutant of B. subtilis 1A382, RB50::[pRF8]<sub>60</sub>(Ade<sup>+</sup>), is produced that is deregulated for riboflavin and purine production and has the rib operon amplified within its chromosome. This mutant is able to produce greater than 5 g/l of riboflavin after 48 hours of fermentation in a 14-liter vessel. Other bacteria are described in which riboflavin production is increased to over 10 g/l under similar conditions.

This invention generally features the production of large quantities (over 10 g/l) of riboflavin by construction of various bacterial strains and growth of those bacterial strains within a medium and under conditions suitable for production of the riboflavin. In a first aspect, the invention features a recombinant bacterium which includes at least one copy of an exogenously introduced nucleic acid within its chromosome. This nucleic acid encodes one or more riboflavin biosynthetic proteins, is heritable, and capable of expression by the bacterium such that riboflavin biosynthesis by the bacterium is increased relative to a bacterium lacking such a sequence.

By "recombinant bacterium" is meant a bacterium which contains one or more nucleic acid sequences, from the same or another organism, at a site at which those sequences do not naturally occur, and/or in a copy number in which they do not naturally occur. Thus, the term includes those bacteria in which two copies of a nucleic acid sequence, e.g., a gene or an operon, are provided at a site which normally includes only one copy of the sequence. It also includes bacteria in which one or more copies of a nucleic acid sequence are introduced at a site which does not normally include that sequence. Such recombinant bacteria are constructed by standard recombinant DNA technology.

By "exogenously introduced" is meant that the nucleic acid is introduced into the chromosome from a source outside of that chromosome by any standard technique, including recombinant DNA technology, transformation, and transfection. It also includes the progeny of such bacteria, for example, those bacteria produced by cellular division of an originally constructed, transformed, or transfected bacterium.

By "riboflavin biosynthetic proteins" is meant to include those peptides, polypeptides or proteins which are directly involved in the synthesis of riboflavin from guanosine triphosphate. These proteins may be identical to those which naturally occur within a bacterium and are involved in the synthesis of riboflavin within that bacterium. Alternatively, they may be modifications of such proteins, for example, they may contain modifications which do not significantly affect the biological activity of the protein. For example, the natural protein may be modified by introducing or substituting one or more amino acids, preferably by conservative amino acid substitution, or by removing nonessential regions of the protein. Such modifications are readily performed by standard techniques.

In some embodiments, the bacterium contains two or more copies of the nucleic acid sequence; and the nucleic acid encoding one or more of the riboflavin biosynthetic proteins is present at at least two sites within the chromosome of the bacterium.

By "site" is meant a distinct chromosomal location relative to a wild-type bacterium at which the nucleic acid encoding the biosynthetic proteins is located. For example, such nucleic acid may be located at the naturally occurring site for genes encoding such proteins (i.e., at a rib locus), or it may be located at a site distant from this location. Preferably such distant sites are chosen from regions of chromosomal nucleic acid which are not essential to the recombinant bacterium, such as regions which encode proteins which are not essential to production of riboflavin. Examples of such regions include those which encode certain extracellular enzymes such as proteases. Insertion at such sites does not interfere with a desirable quality or trait. Any site is suitable as long as the functioning of the bacterium, with regard to riboflavin production, is not substantially affected.

In other embodiments, the nucleic acid is present in a plurality of copies at one or more of the sites; and the nucleic acid is present at at least three sites within the chromosome. By introducing the nucleic acid at different sites, the total number of copies of the nucleic acid within the chromosome can be increased. Increasing the copy number, increases the amount of riboflavin production.

Generally the riboflavin biosynthetic proteins are encoded by one or more rib genes (e.g., an inactivation of which creates a riboflavin auxotroph), preferably at least five distinct rib genes identifiable from the nucleotide sequence provided in Fig. 3. Preferably, at least five copies of such genes are provided. By "rib genes" is meant those genes or portions of genes which encode proteins which occur naturally within an organism, or perform a similar function to such proteins, which are involved in the biosynthetic conversion of guanosine triphosphate to riboflavin within a bacterium.

In a related aspect, the invention features a recombinant bacterium which includes nucleic acid encoding one or more riboflavin biosynthetic proteins, e.g., the gene products identified as ORF1 and ORF6 in Fig. 4, the expression of at least one of which is controlled by a transcription element not naturally associated with the nucleic acid. Alternatively, the recombinant bacterium includes one or more rib genes or transcription units the expression of which is controlled by a transcription element not naturally associated with that rib gene.

By "transcription element" is meant to include any nucleic acid which effects (i.e., turns on) the transcription of nucleic acid downstream from that transcription element. Examples of such elements include promoters and operators. Such transcription elements are not naturally associated with the nucleic acid, for example, they may be heterologous transcription elements. That is, they may be isolated from other species or genera of bacteria or other organisms. Alternatively, the transcription element may be one naturally present in the bacterium but not normally associated with a rib gene to which it is now transcriptionally linked. Such elements do not include those which are naturally associated with a rib gene.

In other embodiments, the recombinant bacterium includes at least three (or at least five) rib genes and the expression of all three rib genes is controlled by a transcription element not naturally associated with those rib genes; at least two transcription elements are provided; the rib genes are provided within the chromosome of the recombinant bacterium; the recombinant bacterium is deregulated for riboflavin gene expression; and the transcription element is a promoter. For example, the promoter is a constitutive, growth-regulated, or inducible promoter, such as one associated with the SPO1 phage, and/or veg, amy, and sacQ-sensitive promoters, e.g., apr.

By "deregulated" is meant that the level of riboflavin production is greater than that observed in a bacterium with natural riboflavin regulatory systems (i.e., a wild type bacterium). Examples of such deregulated bacteria include those which are resistant to various purine analogs or antagonists whereby such analogs or antagonists can e.g. be selected from the group consisting of 8-azaguanine, psicofuranine, decoyinine, 8-azaguanine, sulfaguanine, 6-thioguanine, and methionine sulfoxide, and/or riboflavin analogs e.g. roseoflavin. Preferred deregulated bacteria are resistant to at least one of the following purine analogs or antagonists 8-azaguanine or decoyinine and roseoflavin.

In other specific embodiments, at least one of the rib genes includes a ribosome binding site not naturally associated with the rib gene; the rib genes are present at two sites within the chromosome; and the rib genes are present in a plurality of copies within the chromosome. In more preferred embodiments, the rib genes are Bacillus rib genes, for example ORF3 and ORF4 shown in Fig. 4, and the transcription element is located in a region 5 -upstream of ORF3 or ORF5; and the rib genes are chosen from a  $\beta$ -riboflavin synthase-encoding gene, ORF2, ORF3, ORF4, and ORF5; and the bacterium belongs to a species of Escherichia, e.g., E. coli, Bacillus, e.g., B. subtilis, Klebsiella, or Cornyebacterium.

In another related aspect, the invention features nucleic acid, which includes five or more rib genes, the expression of which is controlled by a transcription element not naturally associated with that rib gene.

In a further aspect, the invention relates to a vector comprising first a nucleic acid sequence of bacterial especially Bacillus whereby B. subtilis is preferred or E. coli or yeast origin coding for one or more, whereby at least five is preferred, riboflavin biosynthetic proteins and second one or more whereby one or two is preferred transcription elements not naturally associated with this nucleic acid sequence. Such vectors are preferred whereby the transcription elements are as already specified above. Further preferred are vectors as specified in the Examples e.g. pRF50, pRF69 pRF70, pRF71, pRF78, pRF81 and/or pRF89.

Another aspect of the invention relates to a recombinant bacterium comprising a bacterium which has been transformed by a vector as specified above whereby at least one copy, preferably a plurality of copies of said nucleic acid sequence including said transcription elements has been introduced at one or more sites with in its chromosome and said nucleic acid sequence including said transcription elements is heritable and capable of expression by the bacterium, such that riboflavin biosynthesis by the bacterium is increased relative to a bacterium lacking such nucleic acid sequence including said transcription elements. Introduction at one or two such sites is preferred. Furthermore recombinant bacteria are preferred whereby the bacterium which is transformed is already deregulated for riboflavin gene expression. Of such deregulated bacteria E. coli or Bacillus especially B. subtilis strains are preferred whereby B. subtilis strains RB50 and RB58 are especially preferred.

The invention relates also to a process for the preparation of a recombinant bacterium whereby a bacterium especially as already specified above is transformed by a vector comprising a nucleic acid sequence of bacterial or yeast origin coding for one or more riboflavin biosynthetic proteins or a vector as already specified above whereby at least one copy, preferably a plurality of copies of said nucleic acid sequence optionally including said transcription elements is introduced at one or more, preferably one or two sites with in its chromosome and said nucleic acid sequence optionally including said transcription elements is heritable and capable of expression by the bacterium, such that riboflavin biosynthesis by the bacterium is increased relative to a bacterium lacking such nucleic acid sequence optionally including said transcription elements.

The invention relates also to bacteria as specified herein wherein one or more transcription elements naturally associated with the rib biosynthetic genes in the chromosome of such bacteria have been replaced by transcription elements as specified. Such replacements can be made on the basis of the description of the present invention and by methods known in the art.

In another aspect, the invention features a method or process for production of riboflavin. The method includes growing cells especially recombinant bacteria as specifically mentioned herein under suitable growth conditions. Such suitable growth conditions are characterized by limiting the availability of a component of the growth medium and/or feed medium in such a way that aerobic conditions for the growth of said recombinant bacterium are maintained. Such conditions can be also characterized e.g. by maintaining a level of dissolved oxygen at a concentration between about 5% to 30%. The man skilled in the art is familiar with the fact that such levels of dissolved oxygen can vary dependent on the specific technical equipment used for growing said recombinant bacteria and for measuring said dissolved oxygen concentration. Under anaerobic conditions the synthesis of riboflavin is reduced. In some embodiments, the limiting component is chosen from a carbon source, nitrogen source, or a component required by the cells (e.g., in the feed medium). For example, if the cells are auxotrophic, for example, for methionine, a limiting level of methionine may be provided in the growth medium. In another example, such component could be glucose or a carbonic acid, e.g. a citric acid cycle acid, such as citric acid or succinic acid, or an amino acid.

In a related aspect, the invention features another method for increasing production of riboflavin by a bacterium. In this method, the strain of bacterium used is deregulated for riboflavin production. More than one copy of a nucleic sequence encoding one or more riboflavin biosynthetic proteins is introduced into the chromosomal DNA of this bacterium. Preferably the bacterium used in this method is selected from one of those described above. The invention relates also to a method for the production of riboflavin by growing a recombinant bacterium obtained by a process as specified above under conditions as also specified above.

In other aspects of the invention, purified nucleic acid and the recombinant polypeptide product of such nucleic acid is provided. Generally, the purified nucleic acid consists essentially of all or a portion of the rib operon, for example, the specific open reading frames shown in Fig. 3. Such purified nucleic acid may  $\overline{be}$  provided within a vector such as a plasmid, phage, or cosmid, or may be integrated within the chromosome of a bacterium. This nucleic acid is separated from nucleic acid with which it is naturally linked. For example, 6.5 kb of the nucleic acid encoding the whole rib operon may be inserted within a Bacillus subtilis chromosome at a site distant from that site in which the 6.5 kb DNA is normally present. By recombinant polypeptide is meant biologically active protein free of extraneous polypeptide (i.e., not fused to a heterologous polypeptide) having an enzymatic activity equivalent to such a naturally produced polypeptide.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

The drawings will first briefly be described:

Figure 1. The riboflavin biosynthetic pathway, modified from Keller et al., Biochem. 27:1117 (1988). The corresponding intermediates shown are those produced by E. coli (which are presumably the same as those produced by B. subtilis): structure 1, guanosine triphosphate (GTP); structure 2, 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone-5´-phosphate; structure 3, 5-amino-6-(ribosylamino)-2,4(1H,3H)-pyrimidinedione-5´-phosphate; structure 4, 5-amino-6-(ribitylamino)-2,4(1H,3H)-pyrimidinedione-5´-phosphate; structure 5, 6,7-dimethyl-8-ribityllumazine; structure 6, riboflavin. The biosynthetic enzymes indicated are those encoded by B. subtili is (GTP cyclohydrolase,  $\alpha$  and  $\beta$  subunits of riboflavin synthase) or those proposed to be encoded by B. subtilis (a rib-specific deaminase, and a rib-specific reductase). Figure 2. Schematic representation of purine biosynthesis. The purine biosynthetic pathway, including the portion responsible for riboflavin biosynthesis, is depicted. The individual enzymes of the pathway are identified by their gene symbols (E. coli nomenclature). Abbreviations are as follows: PRPP, phosphoribosylpyrophosphate; GAR, glycinamide ribonucleotide; pur, GAR formyltransferase; PRA, phosphoribosylamine; purA, adenylosuccinate synthetase; purB, adenylosuccinate synthetase; FGAR,

45

50

55

formylglycinamide ribonucleotide; SAICAR, aminoimidazolesuccinocarboxamide ribonucleotide; purC, SAICAR synthetase; FGAM, formylglycinamidine ribonucleotide; purD, GAR synthetase; AIR, aminoimidazole ribonucleotide; purE, AIR carboxylase; CAIR, carboxyaminoimidazole ribonucleotide; purF, PRPP amidotransferase; AICAR, aminoimidazolecarboxamide ribonucleotide; purH, AICAR formyltransferase; purJ, inosine monophosphate (IMP) cyclohydrolase; FAICAR, formamidoimidazolecarboxamide ribonucleotide; purL, FGAR amidotransferase: guaA, guanosine monophosphate (GMP) synthetase; purM, AIR synthetase; guaB, IMP dehydrogenase.

Figure 3. The complete nucleotide and deduced amino acid sequences of the B. subtilis rib operon. The nucleotide sequence was determined by dideoxy sequencing of M13 clones. The deduced amino acid sequence is indicated by the one letter code (Lehninger, Biochemistry. 2d Ed., Worth Publishers, Inc., New York, p. 72).

Figure 4. A schematic representation of the rib gene cluster. The top diagram is the restriction endonuclease map of the cloned 10 kb EcoRi DNA fragment in plasmid pRF2, containing the B. subtilis rib operon. Homology to the 54-mer probe specific for the riboflavin synthase gene is depicted by the solid bar. The hatched box depicts Rib cloned DNA, while the thin black line represents pBR322 DNA. The bottom diagram is based on the complete nucleotide sequence of the 6.0 kb fragment to which the rib operon was localized. Open reading frames are depicted by open boxes, with arrows indicating the direction of transcription, and closed boxes indicating the putative ribosome binding sites. Probable  $\sigma^{\rm A}$  promoter regions are shown. Tentatively identified rho-independent transcription termination sites are indicated by a "hairpin" symbol. Not all restiction sites are indicated.

Figure 5. Strain lineage of RB50. The lineage of the riboflavin overproducing strain of B. subtilis, RB50, is depicted. The various parent strains were exposed to riboflavin and purine analogs to select appropriate mutations.

Figure 6. Origins of rib $^{\dagger}$ s recombinant plasmids. A schematic diagram of the production of the rib operon-containing recombinant plasmids pRF1, pRF2, pRF3, pRF6 and pRF7 is presented. A library of size-selected, 9-11 kb fragments of B. subtilis DNA was used to produce a gene library in E. coli plasmid vectors. Clones were selected by hybridization to the 54-mer probe specific for the  $\beta$  subunit of the riboflavin synthase gene.

Figure 7. The strain lineage of B. subtilis RB53::[pRFB]<sub>90</sub>. Plasmid pRFB was integrated into the chromosome of the intermediate strain RB52 and amplified; the resulting strain was exposed to the purine analog azaguanine.

Figure 8. Identification of regions essential for riboflavin biosynthesis using insertions and deletions. A diagram is presented of the 10 kb cloned EcoRl DNA fragment with the regions essential for riboflavin biosynthesis indicated. Insertions and deletions at the indicated restriction sites enabled the localization of the rib operon. Not all restriction sites are indicated.

Figure 9. Hairpin-loop structures of the possible rho-independent transcription termination sites. Their locations in the nucleotide sequence of Figure 3 are shown below each structure. Also presented are their free energies of formation, determined according to Tinoco et al. (Nature (London) New Biology 246:40 (1973)).

Figure 10. Structure of various plasmid derivatives used in S-30 in vitro coupled transcriptiontranslation reactions. A schematic diagram is shown of the rib operon regions contained in the plasmid derivatives used in the S-30 reactions, as well as the open reading frames predicted to be expressed.

Figure 11. Comparison of riboflavin production curves. Riboflavin production curves for various fermentation protocols are shown. Open squares: RBF-14 using RB50::[pRF8]<sub>60</sub> (Ade<sup>-</sup>). Closed squares: RBF-22 using RB50::[pRF8]<sub>60</sub> (Ade<sup>-</sup>). Closed circles: RBF-29 using RB50::[pRF8]<sub>60</sub> (Ade<sup>-</sup>).

Figure 12. Construction of pRF40.

Figure 13. Construction of pRF50.

Figures 14, 15, and 16. Structure of various vectors.

Figure 17. 55-mer used in plasmid construction.

Figure 18. Various oligonucleotides used in vector construction.

In the practice of the present invention, host bacterial strains are derived that contain one or more mutations in genes of the riboflavin biosynthetic pathway or in the biosynthetic pathway of various purines, which mutations lead to riboflavin overproduction. In one embodiment, such mutations lead to riboflavin overproduction by deregulating steps in the riboflavin biosynthetic pathway. In another embodiment, the mutations increase riboflavin production by causing an inhibition in the use in an alternative metabolic pathway of a precursor for riboflavin biosynthesis.

In a specific embodiment, desired mutations in the genetic background of the host bacteria can be

5

10

15

20

25

30

35

induced by exposure to analogs of purine or riboflavin that compete with their authentic counterparts in the metabolic pathways of the host; bacteria that survive such exposure will have mutations that allow them to overproduce the authentic counterpart to the analog, thus "competing out" the purine or riboflavin analog that would otherwise be lethal. The biosynthesis of riboflavin in B. subtilis originates with guanosine triphosphate (Figure 1, structure 1). Guanosine triphosphate (GTP), via guanosine monophosphate (GMP), is a product of the purine biosynthetic pathway (Figure 2). In a preferred embodiment, to obtain a host strain that overproduces riboflavin, one can use classical genetics to both increase the amount of GTP that the cell produces and to deregulate the riboflavin pathway. Purine overproduction in B. subtilis can be achieved by obtaining mutants resistant to purine analogs or antagonists. Examples of some of the purine analogs that can be used include but are not limited to 8-azaguanine (Ishii and Shiio, Agric. Biol. Chem. 36:1511, 1972; Konishi and Shiro, Agric. Biol. Chem. 32:396, 1968), psicofuranine and decoyinine (Matsui et al., Agric. Biol. Chem. 43:1739, 1979; Matsui et al., Agric. Biol. Chem. 43:393, 1979), 8-azaxanthine, sulfaguanine, 6-thioguanine (Debabov, V.G. in The Molecular Biology of the Bacilli vol. 1 Bacillus subtilis, D.A. Dubnau, ed. (Academic Press, New York), pp. 331-370, 1982) and others, and/or the antagonist methionine sulfoxide (Matsui et al., App. Env. Microbiol. 34:337, 1977), and any combination thereof.

The riboflavin pathway can be deregulated by obtaining mutants resistant to a riboflavin analog. An example of a riboflavin analog that can be utilized is roseoflavin (Matsui et al., Agric. Biol. Chem. 46:2003, 1982).

In a specific embodiment of the present invention, bacteria that are mutationally resistant to the analogs azaguanine, decoyinine and roseoflavin, can be used. Specific mutants resistant to each of these compounds are described below. Bacteria with mutations rendering them resistant to other analogs can also be used. It is also deemed within the scope of the present invention to utilize bacteria with different mutations rendering resistance to these same analogs, or different combinations of these mutations, either in combination, with or without, various mutations to other analogs.

If exposure to the analog-alone does not produce resistant mutants at a high enough frequency, various mutagens can be used to increase the frequency of mutation in general and thus increase the number of analog-resistant mutants. As one example, ethyl methyl sulfonate can be used, but other mutagens including but not limited to nitrosoguanidine or UV irradiation can also be used.

Suitable bacterial hosts include all Bacilli species (including in a preferred embodiment B. subtilis), E. coli, and many other gram-positive and gram-negative bacteria. Species which can recognize the promoter sequences of the cloned rib operon to be inserted within their genome are suitable for use. The plasmids described below can be used to introduce rib genes into other bacteria by standard procedure, e.g., transformation. Expression of the inserted rib genes can be determined by spectroscopy as described below, or by observation of the bacteria under UV light, as described below.

In addition to creating mutations by exposure to purine or riboflavin analogs, bacterial strains that already contain mutations that are known to affect their purine or riboflavin biosynthetic pathways can be utilized. For example, the present invention makes use of but is not limited to B. subtilis strain 1A382, which contains the mutation pur-60, making it auxotrophic for adenine. Because this mutation blocks the utilization of the riboflavin precursor inosine monophosphate (IMP) in a metabolic pathway other than riboflavin production, increased amounts of IMP are available for riboflavin biosynthesis, thus increasing riboflavin production. There are many other mutations which can be utilized to potentially increase riboflavin production, including but not limited to guaC3, his and others that are included within the scope of the present invention. The guaC3 mutation prevents the conversion of GMP back into IMP (see Figure 2), thus increasing the amount of riboflavin biosynthetic precursors available.

Suitable mutations affecting riboflavin overproduction can be mapped by various methods known in the art. In a specific embodiment, a mutation can be mapped by complementation of auxotrophic mutants.

The riboflavin biosynthetic genes from various bacteria can be cloned for use in the present invention. Yeast or bacterial cells from species including but not limited to the genus Bacillus, E. coli and many other gram-positive-and-gram-negative bacteria can potentially serve as the nucleic acid source for the molecular cloning of the rib operon. The DNA containing the rib operon may be obtained, by standard procedures known in the art, for example, from a DNA library prepared by cloning chromosomal DNA or fragments thereof, purified from the desired bacterial cell, into a suitable vector for propagation of the gene. (See, for example, Maniatis et al., 1982, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1982, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K., Vol. I, II).

In the molecular cloning of the gene from chromosomal DNA, fragments are generated, some of which will encode the desired rib operon. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNASe in the presence of manganese to fragment the DNA, or the

25

DNA can be physically sheared, as for example, by sonication. The linear DNA fragments can then be separated according to size by standard techniques, including but not limited to agarose and polyacrylamide gel electrophoresis and density gradient centrifugation.

Once the DNA fragments are generated, DNA libraries are prepared using an appropriate cloning and/or expression vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. For E. coli such vectors include, but are not limited to, bacteriophages such as λ derivatives, high-copy plasmids such as pBR322 or pUC plasmids, or low-copy plasmids derived from Pseudomonas plasmid RK2. For Bacillus such vectors include, but are not limited to, bacteriophages such as ρ11 Dean et al., J. Virol. 20: 339, 1976; Kawamura et al., Gene 5:87, 1979) or Δ105 derivatives (lijima et al., Gene 9:115, 1980; Errington, J. Gen. Microbiology 130:2615, 1984; Dhaese et al., Gene 32: 181, 1984; Errington, J. in Bacillus Molecular Biology and Biotechnology Applications, A.T. Ganesan and J.A. Hoch, eds. (Academic Press, New York,), p. 217, 1986), high-copy plasmids such as pUB110 (Ehrlich, Proc. Natl. Acad. Sci. (USA) 74: 1680, 1977) or PBD64, or low-copy plasmids such as pE194 derivatives (Gryczan, T.J. in The Molecular Biology of the Bacilli, D.A. Dubnau, ed. (Academic Press, New York), pp. 307-329, 1982; Horinouchi and Weisblum, J. Bacteriol. 150: 804, 1982). Recombinant molecules can be introduced into host cells via transformation, transfection, protoplasting, infection, electroporation, etc.

Once the DNA libraries are generated, identification of the specific clones harboring recombinant DNA containing the rib operon may be accomplished in a number of ways (as described, for example, in Maniatis et. al., supra). For example, if an amount of the operon or a fragment thereof is available from another bacterial source (e.g., from E. coli) and is sufficiently homologous to the riboflavin biosynthetic genes of Bacillus to hybridize thereto, that DNA can be purified and labeled, and the generated bank of DNA fragments may be screened by nucleic acid hybridization to the labeled probes (Benton, W. and Davis, R., 1977, Science 196:180; Grunstein, M. and Hogness, D., 1975, Proc. Natl. Acad. Sci. U.S.A. 72:3961). Alternatively, sequences comprising open reading frames of the endogenous rib operon, or subsequences thereof comprising about 10, preferably 15 or more nucleotides, may be used as hybridization probes. Such probes can be made synthetically, based on a portion of the nucleic acid or amino acid sequence (examples of which are provided below) of a gene product known to be encoded by the operon ("reverse genetics"). If a purified rib operon-specific probe is unavailable, cloned gene libraries of restriction fragments (from partial Sau3A-digests, for example) can be made in bacteria, especially B. subtilis or E. coli, and the rib operon-containing recombinant clones can be identified by either marker-rescue or complementation of known rib mutations.

In a preferred embodiment, the rib operon of B. subtilis can be isolated for use from an E. coli plasmid library of B. subtilis DNA. In particular, and as described below, the B. subtilis rib operon can be isolated by virtue of its homology to a radiolabelled, synthesized nucleotide probe that is derived from an internal region of a gene product known to be encoded by the operon of B. subtilis. Although a portion of the amino acid sequence for β-riboflavin synthase (Ludwig et al., J. Biol. Chem. 262:1016, 1987) can be the basis for such a probe, with the third nucleotide of each codon estimated from frequency of codon usage, a similar probe based on another region of this protein or another protein from the rib operon can be utilized and would fall within the scope of the present invention. The present invention further enables screening by use of synthetic probes which are derived from the nucleic acid sequence shown in Fig. 3.

Analogous methods to those detailed here can be used to isolate the rib operon of other bacteria, especially other Bacilli or E. coli. In a specific embodiment, such clones can be selected by assay for ability to hybridize to the labeled B. subtilis rib operon or a hybridizable portion thereof. It is well known in the art that starting from an appropriate mRNA preparation, cDNA can be prepared; such cDNA can also be used in accordance with the present invention to prepare vectors for the transformation of appropriate bacteria for riboflavin overproduction.

Once the host cells with recombinant DNA molecules that include the isolated rib operon or a portion thereof are identified, the DNA may be obtained in large quantities. This then permits the rib operon to be manipulated and its nucleotide sequence to be determined using various cloning and sequencing techniques familiar to those knowledgeable in the art.

For example, insertional mutagenesis can be used to locate and characterize the rib operon and genes thereof within a cloned piece of DNA. In a specific embodiment, rib-biosynthetic containing regions can be identified by inserting small cat (chloramphenicol acetyltransferase)-containing restriction fragments into several different restriction enzyme sites of the cloned DNA, and testing each derivative for insertional inactivation of riboflavin biosynthesis in an appropriate host (see below).

The cloned DNA corresponding to the rib operon can be analyzed by methods including but not limited to Southern hybridization (Southern, E.M., 1975, J. Mol. Biol. 98:503-517), Northern hybridization (see e.g.,

Freeman et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80 :4094-4098), restriction endonuclease mapping (Maniatis et al., 1982, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), and DNA sequence analysis. Restriction endonuclease mapping can be used to roughly determine the genetic structure of rib operon. Restriction maps derived by restriction endonuclease cleavage can be confirmed by DNA sequence analysis.

DNA sequence analysis can be performed by any techniques known in the art, including but not limited to the method of Maxam and Gilbert (1980, Meth. Enzymol. 65:499-560), the Sanger dideoxy method (Sanger, F., et al., 1977, Proc. Natl. Acad. Sci. U.S.A. 74:5463), or use of an automated DNA sequenator (e.g., Applied Biosystems, Foster City, CA). As an example, the DNA sequence of the rib operon of B. subtilis is presented in Figure 3.

Once the nucleotide sequence of the rib operon has been determined, putative open reading frames (ORFs) can then be identified along with the deduced amino acid sequence of their encoded product. Actual identification of the encoded product can be carried out, e.g., by performing S-30 coupled in vitro transcription/translation reactions, with various ORFs used as templates. Various mutational derivatives of the ORFs can also be tested for activity in functional assays of the S-30 reaction products, in order to test the function of the encoded products.

In a specific embodiment of the invention relating to the B. subtilis rib operon, and detailed in the examples below, the above-described methods were used to determine that B. subtilis riboflavin biosynthesis is controlled by a single operon of approximately 4.2 kb containing five biosynthetic genes: the  $\beta$  subunit of riboflavin synthase and ORFs designated 2, 3, 4, and 5 (see Figure 4). ORFs 2, 3, 4, and 5 were subsequently shown to encode proteins with molecular weights of about 15 kd, 47 kd, 26 kd, and 44 kd, respectively. As described below, ORF 5 was shown to encode a putative rib-specific deaminase that catalyzes the reduction of a deaminated pyrimidine to a ribitylamino-linkage in an early step in riboflavin biosynthesis. Our data also indicated that ORF 4 encodes the  $\alpha$  subunit of riboflavin synthase and ORF 3 encodes a GTP cyclohydrolase, while ORF 2 possibly encodes a rib-specific reductase. ORF 1 and ORF 6 were found to be outside the primary transcription unit of the rib operon. The primary site for initiation of transcription of the rib operon was determined to be probably the apparent  $\sigma^A$  promoter located 290 bp upstream from the first gene in the operon, ORF 5 (Figure 4, P<sub>1</sub>). The coding regions, promoters and transcription termination sites of the B. subtilis rib operon are shown in Table VI below.

The present invention encompasses the nucleotide and amino acid sequences of the genes of the rib operon, as well as subsequences thereof encoding functionally active peptides, and sequences which are substantially the same as such sequences. A functionally active peptide, as used herein, shall mean a protein or peptide which is capable of catalysing a reaction leading to riboflavin biosynthesis. A functionally active nucleic acid sequence shall mean a sequence capable of regulating riboflavin biosynthesis. A sequence substantially the same as another sequence shall mean a sequence capable of hybridizing to the complementary sequence thereof. In addition, a nucleic acid sequence not naturally controlling the expression of a second nucleic acid sequence shall mean a sequence which does not control the expression of the second sequence in the bacterium from which the second sequence is isolated.

Once the genetic structure of the rib operon is known, it is possible to manipulate the structure for optimal use in the present invention. For example, the rib operon can be engineered to maximize riboflavin production.

Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the inserted sequences. When propagating in bacteria the regulatory sequences of the rib operon itself may be used. In an embodiment in which the entire rib operon, or greater than one gene thereof, is desired to be expressed as a polycistronic message, a prokaryotic host is required. In an embodiment in which a eukaryotic host is to be used, appropriate regulatory sequences (e.g., a promoter) must be placed in the recombinant DNA upstream of each gene/ORF that is desired to be expressed.

Specific initiation signals are also required for efficient translation of inserted protein coding sequences. These signals include the initiation codon (ATG, GTG or TTG) and adjacent sequences, such as the ribosome binding site (RBS). It should be noted that the RBS of a given coding sequence can be manipulated to effect a more efficient expression of that coding sequence at the translational level. In cases where an entire open reading frame of the rib operon, including its own initiation codon and adjacent regulatory sequences, is inserted into the appropriate expression vectors, no additional translational control signals may be needed. However, in cases where only a portion of the coding sequence is inserted, or where the native regulatory signals are not recognized by the host cell, exogenous translational control signals, including the initiation codon, must be provided. The initiation codon must furthermore be in phase

with the reading frame of the protein coding sequences to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic.

In addition, a host cell strain may be chosen which modulates the expression of the rib operon gene(s) or modifies and processes the gene product(s) thereof in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus, expression of the genetically engineered rib operon proteins may be controlled. In one embodiment, the regulatory regions of the operon, such as the promoter and the termination/anti-termination regulatory sequences, can be manipulated or replaced with constitutive or growth-regulated promoters to deregulate the rib operon and thus increase riboflavin production. Furthermore, appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the expressed proteins. Many manipulations are possible and within the scope of the present invention.

In one specific embodiment of the invention, the 5 regulatory sequence of the B. subtilis rib operon can be removed and replaced with one or more of several B. subtilis promoters; such a construction will cause high-level expression of the rib biosynthetic genes. This approach would involve the introduction of new restriction sites within a 20-30 bp region between the end of the transcription terminator and the RBS sequence of the first gene in the operon ORF 5. Such restriction sites can be introduced by either site-directed mutagenesis or by deleting all regulatory sequences upstream from the right-most BgIII (BgIII<sub>R</sub>) site located within the first 30 bp of ORF 5 (see Figures 3 and 4) and inserting at this site a synthetic oligonucleotide that finishes off the 5 end of ORF 5 (including the ribosomal-binding site) and contains new upstream restriction sites. Once these constructions are made, promoter-containing restriction fragments with ends compatible to the new restriction sites can be introduced, causing expression of the rib genes under the control of the new promoter. Both constitutive and growth-regulated B. subtilis promoters can be used, including but not limited to strong promoters from the lytic bacteriophage SPO genes, veg, amy (amylase), and apr (subtilisin).

In another aspect of the invention, rib operon DNA fragments which have transcriptional regulatory activity (e.g., promoters) can be used to regulate the expression of heterologous gene products.

According to the present invention, the rib operon can be introduced into bacteria, including for example, Bacilli and E. coli, where it is expressed. In a preferred embodiment, the bacterial host is one of the mutant hosts described above. In a specific embodiment, the cloned rib operon is integrated into the host chromosomal DNA, where it is replicated and expressed along with host genomic DNA. In a most preferred embodiment, multiple copies of the rib operon are integrated into the host chromosomal DNA, thus providing for amplified expression of the rib operon in the deregulated host. One method in which this may be accomplished is chromosomal insertion of a cat-containing rib operon followed by chloramphenicol amplification of the operon, as detailed in the examples sections infra. One can also use a tet<sup>r</sup> gene, or certain other drug resistance genes that are expressed in Bacillus, with the same technique.

In specific embodiments, integration vectors containing the rib operon fragment can be engineered so as to contain the rib operon on the smallest possible DNA fragment, in an attempt to obtain greater amplification of the vector within the host chromosome. For example, vector DNA sequences may be deleted, and/or nonessential DNA flanking the rib operon can be deleted.

In general, bacteria that are prototrophic for riboflavin survive on minimal medium in the absence of riboflavin. Production of riboflavin can be detected and quantified by various methods. In a preferred embodiment, overproduction of riboflavin is readily observed when overproducing bacteria are exposed to UV light at 366 nm, as described infra, producing an observable, yellow fluorescence. For example, many of the engineered plasmids of the present invention are produced in E. coli. For some of these plasmids, overproduction of riboflavin has been confirmed by this method. The amount of riboflavin produced can be quantitated, e.g., with reverse-phase high performance liquid chromatography (HPLC). Cell-free supernatants from bacteria can be fractionated over an HPLC column, as described infra, and monitored for riboflavin at 254 nm. By extrapolation from a standard curve, the concentration of riboflavin can be determined by the area of the peak on the chromatogram.

Riboflavin can also be quantitated by fluorescence spectrophotometry. For example, samples containing riboflavin can be read in a fluorescence spectrophometer set at an emission wavelength of 525 nm and an excitation wavelength of 450 nm.

In addition, other methods known in the art are available to detect or quantitate riboflavin based on its physical and biological properties.

Riboflavin overproducing bacteria can be grown in vessels ranging from shake flasks to large "batch" fermentors, by methods known in the art (see below). In a preferred embodiment, nutrient feed can be manipulated to maximize riboflavin production at the minimum cost by varying the nutrients in the medium.

In a specific embodiment, amplified <u>rib</u> -containing genes can be maintained at high-copy number in the bacterial chromosome by the inclusion of about 60 µg/ml chloramphenicol in the inoculum seed strain (but not necessarily in the fermentor). Chemap 14-liter fermentors can be used at 1000 rpm with a head pressure of 0.6 atmospheres. To achieve a high cell density, various carbon sources such as, for example, glucose, sucrose, citric acid cycle acids, maltose or starch, and various nitrogen sources such as yeast extract, com steep liquor, ammonia and/or protein hydrolysates can be used. It is, however, essential to add parts of these media ingredients in such a way that limitations unfavorable to riboflavin production (e.g., oxygen starvation resulting from too high a carbon source concentration) are avoided. Fermentation media and conditions suitable for use are detailed below.

10

## Examples

15

## Example 1: Riboflavin-Overproducing B. subtilis Mutants

We describe in the examples herein the production of strains of Bacillus subtilis which overproduce riboflavin. In order to accomplish this, we used classical genetics, genetic engineering, and fermentation. Classical genetics with selection using purine and riboflavin analogs was used to deregulate the pathways for purine (riboflavin precursor) and riboflavin biosynthesis. Riboflavin production was increased further by cloning and engineering the genes of the riboflavin biosynthetic pathway (the rib operon), allowing for constitutive, high-level production of rate-limiting biosynthetic enzyme(s).

The biosynthesis of riboflavin in B. subtilis originates with GTP (Figure 1). To obtain a host that overproduces riboflavin we used classical genetics to both increase the amount of GTP that the cell produces and to deregulate the riboflavin pathway. Purine overproduction in B. subtilis can be achieved by obtaining mutants resistant to purine analogs such as azaguanine and decoyinine, and other antagonists such as methionine sulfoxide (see e.g., Ishii and Shiio, Agric. Biol. Chem. 36 (9):1511-1522, 1972; Matsui et al., Agric. Biol. Chem. 43 (8):1739-1744, 1979). The riboflavin pathway can be deregulated by obtaining mutants resistant to the riboflavin analog roseoflavin (Matsui et al., Agric. Biol. Chem. 46(8):2003-2008, 1982). Roseoflavin-resistant strains were selected from several strains which had been previously mutagenized and which were resistant to several purine analogs. Described below are the methods used to produce a strain (RB50) which overproduces riboflavin.

35

## 8-Azaguanine-Resistant Mutants

B. subtilis is effectively killed by the purine analogue 8-azaguanine (Sigma Chemical Co., St. Louis, MO) at a concentration of 500  $\mu$ g/ml, and resistant mutants appear spontaneously at a frequency of less than 1 in  $10^8$ . Ethyl methyl sulfonate (EMS; Sigma) at 30  $\mu$ g/ml was used as a mutagen to increase the frequency of azaguanine-resistant (Agr) mutations. Mutagenesis was performed on cells from B. subtilis strain 168 by standard procedures (Miller, 1972, Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). After plating 4 x  $10^6$  mutagenized cells on minimal medium (Sloma et al., J. Bact. 170:5557, 1988) containing 500  $\mu$ g/ml azaguanine and restreaking for single colonies, 35 Agr colonies resulted. One mutant, RB11 (Agr -11), was used in the construction of RB50.

## Decoyinine-Resistant Mutants

Decoyinine-resistant (Dc<sup>r</sup>) mutations were obtained spontaneously at a frequency of 1 in 10<sup>6</sup> or after EMS mutagenesis at 1 in 10<sup>5</sup> by plating cells on minimal medium containing 100 μg/ml of decoyinine (Upjohn Co., Kalamazoo, Ml). A Dc<sup>r</sup> mutant of RB11 was obtained by mutagenesis with EMS as described above. One Dc<sup>r</sup> colony, RB15 (Ag<sup>r</sup> -11, Dc<sup>r</sup> -15), was used in the construction of RB50.

55

50

## Transfer of the Ag and Dc Mutations

These purine analog-resistant mutations were transferred to a different strain background in order to

isolate them from any unwanted EMS-induced mutations and to verify that the Agr and Dcr mutations were due to single loci. Since part of the "carbon flow" from inosine monophosphate (IMP), a riboflavin precursor, is also used for adenine nucleotide biosynthesis, a host strain was selected that was blocked in the adenosine monophosphate (AMP) pathway via the mutation pur-60, allowing more carbon material to "flow" from IMP to the guanine nucleotide precursors of riboflavin (Figure 2). B. subtilis strain 1A382 (hisH2, trpC2, pur-60) was made competent (Sloma et al., J. Bact. 170 :5557 (1988)) and transformed (by the method of Gryczan et al., J. Bact . 134:318 (1978)) with total DNA prepared from the Agr/Dcr mutant RB15. The Trp -(tryptophan) revertant colonies were selected, with 3.3% (10/300) of those also being Dcr and 2.3% (7/300) Agr. This result was not unexpected since, due to "congression" (transformation of a second unlinked marker), a number of the Trp colonies should also be resistant to decoyinine or azaguanine.

One Dcr colony, RB36 (his H2, pur-60, Dcr-15), one Agr colony, RB40 (his H2, pur-60, Agr-11), and one Dc<sup>r</sup>/Ag<sup>r</sup> colony (which was also found to be his<sup>+</sup>), RB39 (pur-60, Ag<sup>r</sup>-11, Dc<sup>r</sup>-15), were all selected for further study.

## Methionine Sulfoxide-Resistant Mutants

Selection using high levels of methionine sulfoxide (MS; 10 mg/ml, Sigma) resulted in spontaneous mutants appearing at a sufficiently high frequency that mutagenesis with EMS was not necessary. The Agr/Dcr mutant, RB39, was streaked onto minimal medium containing 10 µg/ml MS. Resistant colonies were obtained and were restreaked for single resistant colonies. One strain, RB46 (pur-60, Agr-11, Dcr-15, MSr-46) was selected for further study.

#### Roseoflavin Resistant Mutants

15

Although many of these Agr, Dcr and MSr mutants were likely to be overproducing GTP, none of them produced levels of riboflavin detectable on plates. In order to deregulate the riboflavin biosynthetic pathway, conditions were determined to select for resistance to the riboflavin analog roseoflavin (Toronto Research Chemical). Maximum killing of cells occurred at 100 µg/ml of roseoflavin in minimal or complete medium; increasing the concentration did not result in any additional killing. Mutations to roseoflavin resistance (RoF') spontaneously occurred at a sufficiently high rate (approximately 5 x 10<sup>-5</sup>) such that mutagenesis with EMS or other chemicals was not necessary.

Approximately 1000 RoFr colonies were obtained from each of the strains described above, 1A382, RB36, RB39, RB40 and RB46. RoFr mutants from all of these strains showed a low level of fluorescence on minimal media plates when exposed to long-wave UV light (366 nm), indicating some riboflavin production. One of the RoFr colonies obtained from RB46, RB46Y (pur-60, Agr-11, Dcr-15, MSr-46, RoFr-46), when grown on minimal medium, produced 14 mg/l of riboflavin as determined by HPLC (described above).

Of all the strains treated, only RB39 and RB46 produced a significantly different phenotype when RoFr colonies were selected. Approximately 0.5% to 1.0% of the RoFr colonies of either RB39 or RB46 produced an intensely fluorescent, yellow colony. Of these colonies, RB51 (pu r -60, Agr-11, Dcr-15, RoFr-51), arising from RB39, and RB50 (pur-60, Agr-11, Dcr-15, MSr-46, RoFr-50), arising from RB46, produced a stable, fluorescent-yellow phenotype which correlated with a higher level of riboflavin production, as determined by HPLC. When grown in minimal medium, both RB50 and RB51 produced higher levels of riboflavin in their supernatants than the other RoFr strains, about 40 mg/l and 30 mg/l, respectively. The lineage of RB50 is depicted in Figure 5.

Because intensely fluorescent (and thus riboflavin overproducing) colonies could be obtained in non-MSr strains such as RB51, it appeared that this mutation in general might not be contributing significantly to the higher production phenotype. Both of the other mutations, Agr and Dcr (Agr-11 and Dcr-15 in RB39), appear to be necessary to produce high levels of riboflavin since no intensely fluorescent RoFr colonies could be found in strains containing only the Agr-11 (from RB40) or Dcr-15 (from RB36) mutation alone.

## guaC Mutations

Another possibly important mutation for achieving overproduction of GTP, and thus riboflavin, is guac3, which prevents the conversion of GMP back into IMP (see Figure 2). To construct a strain containing guac3 that overproduces riboflavin, competent B. subtilis strain 62121 cells (guaC3, trpC2, metC7) (Endo et al., J.

Bact. <u>15</u>: 169, 1983) were transformed with RB50 DNA and selected for Dc<sup>r</sup> on plates containing 100 µg/ml of decoyinine. Thousands of Dc<sup>r</sup> colonies resulted. Of 200 colonies which were patched onto Dc<sup>r</sup> plates, one was found that exhibited the riboflavin overproduction phenotype (based on UV fluorescence), and was RoF<sup>r</sup>. This colony was designated RB52 (guaC3, trpC2, metC7, Dc<sup>r</sup>-15, RoF<sup>r</sup>-50) and was reserved for subsequent study.

## Other Analog-Resistant Mutants

10

30

35

Finally, because mutants resistant to several additional purine analogs also have been reported to be altered in purine metabolism, such mutations were assayed in order to investigate their effect on riboflavin-overproducing strains. It was determined that 500 g/ml of 8-azaxanthine, 1 mg/ml of 6-thioguanine, or 2 mg/ml of sulfaguanidine (Sigma) effectively kills wild-type B. subtilis. The azaguanine-resistant, riboflavin-overproducing strains RB50::[pRF8]<sub>90</sub> and RB53::[pRFB]<sub>90</sub> (see below) were found to be already resistant to azaxanthine. Although separate azaguanine-and azaxanthine-resistant mutations with different properties have been described previously, in this case the Ag<sup>r</sup>-11 and Ag<sup>r</sup>-53 mutations appear to also convey azaxanthine resistance.

## HPLC Analysis of riboflavin in crude supernatants of B. subtilis

Accumulation of riboflavin in B. subtilis cultures was quantitated by reverse-phase HPLC. Riboflavin standards (Sigma Chemical Co., St. Louis, MO) or cell-free supernatants from strains to be tested were fractionated over a 4.6 mm  $\times$  250 mm Vydac C<sub>18</sub> column equilibrated with 1% ammonium acetate (pH 6.0). At injection, the column was developed with a linear gradient of methanol and monitored for riboflavin at 254 nm. Authentic riboflavin (i.e. riboflavin "standard") elutes at the mid-point of the gradient.

## Example 2: Cloning B. subtilis rib operon

Our general strategy to isolate a restriction fragment containing the rib operon was to screen a "mini" E. coli plasmid library of B. subtilis DNA by hybridization with a synthetic oligonucleotide probe, the DNA sequence of which was partially derived from the published amino acid sequence for the  $\beta$  subunit of riboflavin synthase (Ludwig et al., J. Biol. Chem. 262:1016, 1987). A summary of the protocol is presented in Figure 6.

A synthetic, 54-base "guess-a-mer" oligonucleotide probe was used for this screening based on amino acids 84-102 of the 240 amino acid riboflavin synthase protein, sequenced by Ludwig et al. (J. Biol. Chem. 262:1016-1021, 1987). The third nucleotide of each codon in the probe was chosen according to estimates made of the most frequent codon usage of B. subtilis, based upon, for example, some of the sequences available in GenBank® (Los Alamos Nat. Lab, Los Alamos, NM). The probe consisted of the following sequence:

# 5'-GGAGCTACAACACATTATGATTATGTTTGCAATGAAGCTGCTAAAGGAATTGCT-3'.

To test the specificity of the probe, the <sup>32</sup>P-labelled 54-mer DNA was hybridized to nylon filters containing EcoRl-digested chromosomal DNA (Southern, J. Mol. Biol . 98:503, 1975) isolated from wild-type and the mutant B. subtilis strains. The probe strongly hybridized to a single 9-10 kb fragment of EcoRl-digested B. subtilis ( ribs met -) DNA, which is in good agreement with the predicted size of the rib-containing fragment (Osina et al., FEBS. Lett. 196:75, 1986). A labelled fragment of the identical size was detected when the probe was hybridized to two mutant strains, RB46 (pur-60, Agr-11, Dcr-15, MSr-46) and RB50 (pur-60, Agr-11, Dcr-15, MSr-46, RoFr-50), the latter being a riboflavin overproducer. These hybridization experiments were repeated using HindIll-cut chromosomal DNA, which resulted in the probe identifying a smaller, single fragment of approximately 1.8 kb; this latter result was useful in determining the general location of the rib biosynthetic operon within the cloned DNA.

# Isolation of Plasmids pRF1, pRF2 and pRF3, Containing Wild-type rib Biosynthetic Genes

A "mini" gene library of 9-11 kb EcoRI fragments from B. subtilis strain 168 (rib\*) DNA was prepared using pRK290, a low-copy number vector derived from the Pseudomonas replicon RK2 (Ditta et al., Plasmid

13:149, 1985). ECoRI fragments (size 9-11 kb) of B. subtilis (rib  $^{+}$  met  $^{-}$ ) DNA were isolated by sucrose (10- $\overline{40}$ %) rate-zonal centrifugation. A four-fold excess of these fragments (0.22  $\mu$ g) was ligated to EcoRI-cut pRK290 (0.26  $\mu$ g), that had been dephosphorylated with calf intestinal alkaline phosphatase (CIAP), at a total DNA concentration of 10  $\mu$ g/ml. Approximately 10 ng of ligated DNA was transformed into E. coli DH5 (F-, endAl, hsdR11 [ $r_k$ -,  $m_k$ +], supE<sub>44</sub>, thi-1,  $\lambda$ -, recAl, gyrA96, relAl), resulting in tetracycline-resistant (Tc') colonies at a frequency of 7.7 x 10<sup>4</sup>/ $\mu$ g of DNA. To determine the fraction of transformants containing insert DNA of 9-11 kb, plasmid mini-lysates were prepared from several Tc' transformants, and their DNA was analyzed by restriction enzyme digestion. About 40% of the Tc' transformants were found to contain single EcoRI-generated inserts of 9-11 kb.

Approximately 1140 of the Tc<sup>r</sup> colonies were screened with the <sup>32</sup>P-labelled 54-mer probe specific for the riboflavin synthase gene. One colony gave a positive signal. Plasmid DNA, designated pRF1, was isolated from this clone and tested for Rib<sup>+</sup>-marker rescue activity by transforming the DNA into B. subtilis 1A210 that contains the riboflavin-deficient mutation rib-2, and selecting for Rib<sup>+</sup> prototrophic colonies. pRF1 transformed 1A210 to Rib<sup>+</sup> prototrophy at a high frequency. Plasmid DNA from a randomly chosen Tc' transformant failed to rescue this marker.

Restriction enzyme analysis revealed that pRF1 actually contained two EcoRI-fragment inserts, of 10 kb and 11 kb. To determine which fragment contained the rib operon, EcoRI-digested PRF1 was probed with the <sup>32</sup>P-labelled, 54-mer riboflavin synthase probe. The results indicated that only the smaller, 10 kb fragment cross-reacted with the probe. Moreover, when the 10 kb ECoRI fragment was recloned into the EcoRI site of pBR322, recombinant plasmids pRF2 and pRF3 resulted, representing the two possible orientations of insertion. Both plasmids were found to rescue the rib-2 mutation of B. subtilis 1A210 to prototrophy at a high frequency.

## Isolation of Plamsids pRF6 and pRF7 Containing rib Biosynthetic Genes From RoFr-B. subtilis Strain RB50

RB50 is one of the RoF' mutants of B. subtilis, produced as described above, that is deregulated for riboflavin biosynthesis. It has been reported that approximately 80% of RoF' mutations reside within the rib operon at the rib0 locus (Stepanov, et al., Genetika (USSR) 13:490, 1977). Like the wild-type rib operon, rib genes in RB50 were also contained on a 9-10 kb EcoRl fragment; thus this fragment was cloned using the protocol outlined in Figure 6, with pBR322 used as the cloning vector. Size-selected 9-11 kb EcoRl fragments (0.1  $\mu$ g) from RB50 were prepared as before and ligated to a two-fold excess of ends of EcoRl-cut, dephosphorylated pBR322 DNA (0.34  $\mu$ g) at a total DNA concentration of 22  $\mu$ g/ml. Approximately 9 ng of ligated DNA was transformed into E. coli DHS, resulting in ampicillin-resistant (Ap') colonies at a frequency of 3.5 x 10<sup>5</sup>/ $\mu$ g of DNA.

Restriction enzyme analysis of plasmid DNA isolated from a sampling of 12 Apr colonies revealed that 50% contained plasmids with 9-11 kb EcoRI inserts. Approximately 1140 Apr colonies were screened with the <sup>32</sup>P-labelled 54-mer probe specific for the riboflavin synthase gene by colony hybridization. Six colonies gave positive signals. Plasmids pRF6 and pRF7, isolated from two of these six colonies, were identified by restriction enzyme analysis as containing inserts with the same orientation as pRF2 and pRF3, respectively. In addition, both plasmids were able to marker-rescue the rib-2 mutation at high frequencies.

# Example 3: Introducing rib DNA Into B. subtilis

As described supra, the rib operon from both a wild-type strain and a RoFr mutant of B. subtilis were cloned as identical 10 kb EcoRl fragments into the EcoRl site of the E. coli replicon pBR322; the derivation of these recombinant plasmids is schematically diagrammed in Figure 6. To introduce the 10 kb EcoRl fragment containing the rib operon into B. subtilis in multiple copies, and thus further increase riboflavin production, we constructed a plasmid vector which would allow integration into the B. subtilis chromosome. The integrated DNA was amplified by selecting colonies that would grow at high drug concentrations of chloramphenicol.

## Construction of and Transformation with Integrational rib Plasmids pRF4 and pRF8

To construct the integrational vector, the drug-resistance gene chloramphenicol acetyltransferase (cat), which is selectable in B. subtilis, was introduced into pRF2 and pRF6, the pBR322 vectors with the 10 kb

fragment from wild-type or RoFr B. subtilis strains, respectively. The plasmids pRF2 and pRF6 were digested with BamHI, which cuts the plasmids uniquely within the pBR322 sequence, and dephosphorylated with CIAP. The cleaved DNA was ligated to a 1.3 kb BamHI fragment containing the cat gene (Youngman et al., Plasmid 12: 1-9, 1984), and the ligated DNAs then transformed into E. coli DHS cells (Hanahand, J. Mol. Biol. 166: 557, 1983). Approximately 80-90% of the Apr transformants were chloramphenical resistant (Cmr); restriction analysis of the isolated plasmids (Maniatis et al.) confirmed that plasmid DNA from the Cmr colonies contained the 1.3 kb fragment. The plasmid containing the wild-type riboflavin fragment and the cat gene was designated PRF4; the plasmid containing the cloned riboflavin fragment from the RoFr strain was called pRF8. (Since the RoFr mutation was subsequently shown to be outside the rib operon, these plasmids are presumably identical).

The plasmids pRF4 and pRF8 were transformed into four different B. subtilis strains: the riboflavin overproducer RB50 (Agr-11, Dcr-15, MSr-46, RoFr-50), the RB50 parent RB46 (Agr-11, Dcr-15, MSr-46), the RB50 parent 1A382 and IS75, a common laboratory strain. Competent IS75 and 1A382 cells were transformed with pRF4 or pRF8; these same plasmids were introduced into RB46 and RB50 by transformation of protoplasts (Chang and Cohen, Mol. Gen. Genet 168 :111-115, 1979). The pRF4 or pRF8 DNA integrated into each of these four strains was amplified  $\overline{by}$  selecting for colonies that grew at higher chloramphenicol concentrations. In each strain, we were able to obtain colonies that grew in up to 60  $\mu$ g/ml of chloramphenicol.

In addition, RB52 (guaC3, trpC2, metC7 Dc<sup>r</sup>-15, RoF<sup>r</sup>-50), produced by transforming the guaC3 B. subtilis strain 62121 with DNA from RB50, was made competent and transformed with pRF8. The integrated plasmid in one of the many Cm<sup>r</sup> colonies that resulted was amplified using 90 μg/ml of chloramphenicol. The resulting cells, RB52::[pRF8]<sub>50</sub>, were grown to mid-log phase and plated on minimal media containing 500 μg/ml azaguanine. Approximately 20 Ag<sup>r</sup> colonies resulted. One such colony seemed to produce a more intense fluorescence. The lineage of this strain, RB53::[pRF8]<sub>50</sub>, is given in Figure 7.

## Example 4: Riboflavin Overproduction by Strains Containing pRF4 or pRF8

RB50 containing pRF4 or pRF8 displayed the riboflavin overproduction phenotype (yellow and UV-fluorescent colonies). Amplification of the rib DNA in a wild-type strain or the parent strains of RB50 did not yield yellow or UV-fluorescent colonies, a finding that indicates that the RoFr mutation (which deregulates the biosynthesis of riboflavin) is required for chromosomal amplification of wild-type DNA to cause riboflavin overproduction. A series of shake flask fermentations were performed in 25 ml of riboflavin minimal medium (RMM, in Table I) in a 300 ml baffled flask (Bellco) to measure the production of riboflavin from RB50 that contained the integrated and amplified rib operon.

TABLE I

COMPOSITION OF M	EDIA					
RMM	g/l					
Sodium glutamate	2.0					
Casamino acids (Difco)	0.2					
Yeast extract (Difco)	0.2					
KH <sub>2</sub> PO <sub>4</sub>	6.0					
K₂HPO₄	14.0					
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	2.0					
Sodium citrate	1.0					
MgSO <sub>4</sub> .7H <sub>2</sub> O	0.2					
Adenosine	0.05					
(adjusted to pH 7.0 and autoclaved)						
Maltose	15.0					
(added as sterile 20% solution after autoclaving)						

55

50

40

45

The fermentations were run with strains RB46, RB50 and RB50 containing pRF4 amplified by selection for resistance to 30  $\mu$ g/ml of chloramphenicol (RB50::[pRF4]<sub>30</sub>) and 90  $\mu$ g/ml of chloramphenicol (RB50::[pRF4]<sub>90</sub>). At 24 and 48 hours, supernatant samples were removed and measured for riboflavin content by reverse-phase HPLC.

As shown in Table II, RB50::[pRF4]<sub>30</sub> produced 0.3 g/l of riboflavin, and RB50::[pRF4]<sub>90</sub> produced 0.7 g/l of riboflavin, in 48 hours, which is significantly more than that produced by the strains without rib amplification, such as RB46 and RB50.

TABLE II

QUANTITATIVE ANALYSIS OF RIBOFLAVIN-CONTAINING SUPERNATANTS FROM B. SUBTILIS Riboflavin\* **Culture Time** Strain (hours) (g/l) 0.009 **RB46** 24 **RB50** 24 0.02 RB50::[pRF4]30 24 0.1 RB50::[pRF4]90 24 0.4 0.007 **RB46** 48 0.05 **RB50** 48 RB50::[pRF4]30 48 0.3 RB50::[pRF4]90 0.7

The dramatic increase in riboflavin production resulting from amplification of rib genes in the deregulated host argues that information encoded by the cloned DNA is rate-limiting for riboflavin biosynthesis.

#### Example 5: Mapping the RoFr-50 Mutation

The RoFr-50 mutation in RB50 appeared to be critical to the riboflavin-overproduction phenotype. To identify and possibly move the mutation into different strain backgrounds it was necessary to map the location of the RoFr-50 mutation on the B. subtilis chromosome. Since pRF4 and pRF8 gave very similar levels of riboflavin production in all strain backgrounds, it seemed unlikely that the RoFr-50 mutation was located on the cloned 10 kb EcoRI, rib-containing fragment. More likely, the RoFr-50 mutation is an unlinked repressor-type mutation, possibly in ribC, a repressor mutation which has been reported to map in the lysaroD region of the B. subtilis chromosome (Chernik et al., Genetika (USSR) 15:1569, 1979). To determine whether the RoFr-50 mutation was linked or unlinked to the riboflavin operon, competent B. subtilis 1A210 (rib-2) cells were transformed with RB50 DNA, selecting for rib\*. Thousands of rib\* colonies resulted, and 200 colonies were patched onto tryptose blood agar base containing 100 g/ml of roseoflavin. No RoFr colonies resulted, and none of the colonies exhibited the riboflavin overproduction phenotype, confirming that the RoFr-50 mutation is not located in the rib operon.

# Example 6: Locating rib \* Biosynthetic Genes Using CAT Insertional Mutagenesis

Figure 4 contains a restriction map of the rib-containing 10 kb EcoRI fragment of pRF2, prepared according to standard procedures. Restriction enzyme sites for Xbal, Bglll, Sstl, Hpal and Ncol are unique to the insert DNA, whereas Sali and Pstl cut once in the insert and once in the vector; the insert does not contain any BaMHI, Xhol or Nhel restriction sites. Restriction enzyme Hindlll cleaves the insert at multiple sites; the 54-mer probe specific for the riboflavin synthase gene hybridized to an approximately 1.8 kb Hindlll fragment, suggesting that the rib operon must also reside in the general area surrounding the Sall

10

15

20

25

35

<sup>\*</sup> Riboflavin was measured using an HPLC assay.

and left-most Bglll (BglllL) sites.

In general, to determine the boundaries of the rib operon, small cat-containing restriction fragments were used to construct insertions and deletions in the rib -cloned DNA fragment of pRF2. E. coli plasmid pEccl served as the primary source of restriction fragments bearing a cat gene which confers chloramphenicol-resistance in both E. coli and B. subtilis. This plasmid, a derivative of pMI1101 (Youngman et al., Plasmid 12, 1-9, 1984) in which a non-essential region of the plasmid was removed by standard recombinant DNA techniques, contains a 1.3 kb cat-containing fragment flanked by the "polylinkers" of M13mp7, and therefore is capable of generating cat cassettes with either Smal, EcoRI, Sall or BamHI ends. To generate Sstl or Xval-ended fragments containing the cat gene, the 1.3 kb cat-containing BamHI fragment of pEccl was isolated, the ends modified with HindIII linkers, and the modified fragment cloned into the HindIII site within the polylinker region of pIC2OR, generating plasmid pEcc4.

Integrative plasmid derivatives were first constructed in E. coli and then transferred to the rib chromosomal locus of B. subtilis by DNA transformation. This was done by linearizing the plasmid by a restriction enzyme cut outside the cloned DNA insert, transforming competent B. subtilis strain 1A382 or PY79 ( $\beta^c$ , rib ) cells with this cut DNA, and selecting for Cmr. Because the pBR322 replicon is unable to replicate in B. subtilis, and the cat gene is bounded on both sides by sequences homologous to the rib locus, the cat-containing insertion or deletion can only be inserted into the chromosome by a double-crossover recombination event to yield Cmr transformants. To determine whether the insertion or deletion inactivated riboflavin synthesis, Cmr colonies were assessed for growth on minimal medium agar plates with or without the presence of riboflavin (Rib phenotype).

As diagrammed in Figure 8, cat-containing restriction fragments were inserted by ligation into the individual restriction sites for Xbal, Sstl, Sall and Bglll of pRF2, inserted between the pair of Bglll or Ncol sites (generating deletions removing either a 2.0 kb Bglll fragment or a 0.8 kb Ncol fragment) or inserted into single Haelll and EcoRV sites of the approximately 1.8 kb Hindlll fragment that hybridized to the ribspecific DNA probe, according to standard techniques. The results are shown in Table III.

30

. 35

40

45

50

## TABLE III

# CHARACTERIZATION OF INSERTION AND DELETION DERIVATIVES OF rib± DNA

5	Insertion derivative <sup>a</sup>	<u>B. subtilis</u> b <u>Riboflavin Phenotype</u>
10	A( <u>Xba</u> I) r l B( <u>Sst</u> I <sub>L</sub> ) r	+ ND + ND
15	C( <u>Sst</u> I <sub>R</sub> ) r 1 D( <u>Bgl</u> II <sub>L</sub> )	ND  
20	r 1 E( <u>Sal</u> I) r 1 F( <u>Bgl</u> II <sub>R</sub> )	  
25	G(HaeIII) r	  ND +
30	H( <u>Eco</u> RV) r l <u>Deletion derivative</u>	ND
35	Bgl r l Nco	 
40	r 1	+ +
<b>4</b> 5	transcriptional relative to the	l "l" (left) identify the L orientation of the inserted <u>cat</u> gene e restriction map in Figure 8.
	b B. subtilis str hisH2) or PY79 (SP 8	rain 1A382 (rib+, trpC2, pur-60, rib+)

As summarized in Figure 8 and Table III, insertions into the Sall, either BgIll, or the "right most" Sstl (Sstl<sub>R</sub>) sites, or deletion of the 2.0 kb BgIll fragment, all generated Cm<sup>r</sup> colonies that could not produce riboflavin (Rib<sup>-</sup>), indicating that the rib operon was centrally located within the cloned DNA. Significantly, removal of the 0.8 kb Ncol fragment apparently had no effect on riboflavin production (Rib<sup>+</sup>), suggesting that one end of the rib gene cluster was located to the left of the "left most" Ncol (Ncol<sub>L</sub>) site. The other end of the rib operon was initially determined to map within the approximately 1.8 kb HindIII fragment because the two insertions at sites within the fragment, EcoRV and HaelII, as well as sites distal to the fragment, Xbal and Sstl<sub>L</sub>, all generated Cm<sup>r</sup> colonies that produced riboflavin.

## Example 7: Nucleotide Sequence of the rib Operon

Based on the cat-insertional mutagenesis of the cloned 10 kb DNA fragment, the entire rib operon was localized within a 6.0 kb region bounded by the Sstl<sub>L</sub> and Ncol<sub>L</sub> sites.

This 6.0 kb region of pRF2 containing the rib operon and flanking regions was sequenced by the dideoxy method of Sanger et al. (Proc. Natl. Acad. Sci. USA 74:5463, 1977). Briefly, M13 clones for sequencing were prepared either by subcloning specific restriction fragments into M13, by using the exonuclease activity of T4 DNA polymerase to generate a series of overlapping deletions (Dale et al., Plasmid 13:31, 1985), or by "shot-gun" cloning random fragments, from sonicated restriction fragments, into M13. In some cases, the nucleotide sequence across a restriction site juncture of adjacent fragments was also determined by primer extension sequencing. Approximately 5500 bp were sequenced on both strands and analyzed for sequences resembling typical open reading frames with gram positive-bacteria ribosome binding sites, gram-positive promoters and rho-independent transcription termination sites.

Analysis revealed six complete, non-overlapping open reading frames (Figure 3): ORF 2 (124 amino acids), the gene coding for the  $\beta$  subunit of riboflavin synthase (154 amino acids), ORF 3 (398 amino acids), ORF 4 (215 amino acids), ORF 5 (361 amino acids) and ORF 6 (105 amino acids). Each ORF was preceded by a strong Bacillus ribosome binding site (RBS) with calculated thermostability ranging from  $\Delta G$  = -16 to -22 kcal/mol, and all of them were oriented in the same transcriptional direction. In addition, within the coding region of ORF 5, a second RBS site and ATG start codon were identified, potentially encoding a smaller protein of 248 amino acids. However, based on S-30 in vitro coupled transcription/translation reactions (see below), ORF 5 appears to encode only a 361 amino acid protein. Finally, part of another coding region, ORF 1, encoding the last 170 amino acids of a protein and oriented in the opposite direction, was also identified.

Based on the following observations, riboflavin biosynthesis in Bacillus is controlled by a single operon containing 5 genes: the  $\beta$  riboflavin synthase gene, ORF 2, ORF 3, ORF 4, and ORF 5, of which at least four, the  $\beta$ -riboflavin synthase gene, ORF 3, ORF 4 and ORF 5, unambiguously encode biosynthetic enzymes, with the remaining one, ORF 2, possibly encoding a biosynthetic enzyme.

- 1. ORF 3, ORF 4 and ORF 5 overlap restriction enzyme sites where insertion of cat-containing restriction fragments caused inactivation of riboflavin production in B. subtilis (Figures 4 and 8).
- 2. ORF 1 overlaps a restriction enzyme site(s) where insertion of cat-containing restriction fragments did not cause inactivation of riboflavin production in a rib B. subtilis strain (Table III and Figure 8), nor did it cause reduction of riboflavin production in the deregulated, RoFr B. subtilis strain RB52.
- 3. ORF 2 also overlaps a restriction enzyme site, EcoRV, where insertion of cat-containing restriction fragments did not cause inactivation of riboflavin production in a rib B. subtilis strain (Table III and Figure 8). However, such an insertion did cause a detectable reduction of riboflavin production in the deregulated, RoFr B. subtilis strain RB52, indicating that the mutated ORF 2 gene product was partially inactive for riboflavin production. The results suggest that ORF 2 does encode a rib-specific enzyme.
- 4. Two DNA sequences capable of forming stem-loop structures indicative of rho-independent transcriptional termination sites were identified within the intercistronic gaps between ORF 1 and ORF 2 and between ORF 5 and ORF 6 (Figures 4 and 9). Removal of structures between ORF 5 and ORF6 enhances expression of riboflavin. The structures impart riboflavin sensitivity to lacZ-fusion constructs; thus, they can be used to impart such sensitivity to any other gene to which they are fused at the 5'-end upstream of the promoter.
- 5. A DNA sequence, TTGCGT-(17bp)-TATAAT, resembling the promoter recognized by the  $\sigma^A$  (vegetative form) of B. subtilis RNA polymerase was identified approximately 290 bp upsteam from ORF 5, oriented in the same transcriptional direction as ORF 5 (Figure 4). A transcriptional fusion of this promoter (P<sub>1</sub>, on a 1.1 kb BgIII-Ncol restriction fragment) to a promoterless E. coli lacZ gene (P<sub>1</sub>-lacZ) displayed riboflavin-regulated expression of  $\beta$ -galactosidase activity in a rib<sup>+</sup>, B. subtilis strain (62121) and high-level, constitive (unregulated) expression of  $\beta$ -galactosidase activity in a rib<sup>+</sup>, RoF<sup>r</sup> B. subtilis strain (RB52) only when the promoter was oriented in the same transcriptional direction as the gene, as shown in Table IV. Primer extension analysis was used to confirm the start site. Transcriptional and Northern analyses were used to show a polycistronic RNA of 4.2 kb encompasses the entire rib operon.

5

25

30

35

40

45

#### **TABLE IV**

RIBOFLAVIN-REGULATED EXPRESSION OF P1-LacZ TRANSCRIPTIONAL FUSIONS							
	β-Galactosidase Specific Activity (Miller Units)						
Strain (integrated plasmid)	+ Riboflavin (2 µg/ml)	- Riboflavin					
B . subtilis 62121 (P <sub>1</sub> - lac Z <sup>a</sup> )	1.3	4.2					
B . subtilis RB52 (P <sub>1</sub> - lac Z <sup>a</sup> )	31	38					
B . subtilis 62121 (P <sub>1</sub> - lac Z <sup>b</sup> )	<0.1	<0.1					
B . subtilis 62121	<0.1	<0.1					

<sup>&</sup>lt;sup>a</sup> P<sub>1</sub> and lac Z orientated in the same direction

25

30

40

45

50

55

5

10

15

- 6. A second DNA sequence, TTGAAG-(17bp)-TACTAT, resembling a promoter recognized by the  $\sigma^A$  (vegetative form) of B. subtilis RNA polymerase was identified within the 3 end of ORF 4, approximately 295 bp upstream from ORF 3 and oriented in the same transcriptional direction as ORF 3 (Figure 4). Integration into B. subtilis by a Campbell-type recombination event of an E. coli plasmid containing this promoter sequence on a 0.7 kb Sall-BgIII restriction fragment did not cause inactivation of riboflavin production in B. subtilis, results which indicated that this second sequence (P2) has promoter activity and thus may actually control transcription (in addition to the  $\sigma^A$  P1 promoter) of ORF 3, the  $\beta$  subunit riboflavin synthase gene and ORF 2. LacZ fusions and Northern analysis confirmed the existence of this promoter.
- 7. A third DNA sequence, TTGAAT-(18bp)-TAAAAA, possibly resembling the promoter recognized by the  $\sigma^A$  (vegetative form) of B. subtilis RNA polymerase was identified within the intercistronic region between the  $\beta$  subunit of the riboflavin synthase gene and ORF 2, approximately 83 bp upstream of ORF 2 and oriented in the same transcriptional direction (Figure 4). This  $\sigma^A$  promoter,  $P_3$ , may also control transcription of ORF 2, in addition to  $P_1$  and  $P_2$ .
  - 8. In vitro-coupled transcription/translation analysis of S-30 reactions of the cloned DNA confirmed that ORF 2, ORF 3, ORF 4, and ORF 5 all actually encoded proteins of the size predicted from their respective sequences.
  - 9. Three of the five presumed enzymatic steps in riboflavin biosynthesis were assigned to specific coding regions by comparing predicted amino acid sequences or molecular weights of their products to published protein sequences, using GenBank®, or known protein sizes.
    - a. The putative protein encoded by the open reading frame between ORF 2 and ORF 3 almost identically matched the published 154 amino acid sequence of the β subunit for the riboflavin synthase enzyme (Ludwig et al., J. Biol. Chem. 262:1016, 1987). Only one amino acid difference was detected: lysine was substituted for glycine at residue 65. This enzyme is reported to catalyze the formation of 6,7-dimethyl-8-ribityllumazine from 5-amino-6-ribitylamino-2,4(1 H , 3 H )-pyrimidinedione-5'-phosphate (Figure 1, structures 5 and 4, respectively) and 3, 4-dihydroxybutanone-4-phosphate.
    - b. A 39% identity in an 88-amino acid overlap was identified between the putative product of ORF5 and deoxycytidylate deaminase, a 188 amino acid protein encoded by the E. coli bacteriophage T<sub>2</sub> (Maley et al., J. Biol. Chem. 258:8290, 1983). Based on this result, ORF 5 most likely encodes the rib-specific deaminase that catalyzes the formation of 5-amino-6-(ribosylamino)-2,4(1H,3H)-pyrimidinedione-5'-phosphate from 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone-5-phosphate (Figure 1, structures 3 and 2, respectively).

<sup>&</sup>lt;sup>b</sup>  $P_1$  and  $\overline{lac}$  Z orientated in opposite directions Based on these results, this  $\sigma^A$  promoter,  $P_1$ , is a primary promoter for transcription of ORF 5, ORF 4, ORF 3,  $\beta$ -riboflavin synthase gene and ORF 2.

c. The predicted molecular weight of the ORF 4 gene product (26,000 Da) was in good agreement with the molecular weight of the  $\alpha$ -subunit for riboflavin synthase (23,000 Da; Bacher et al., J. Biol. Chem. 255:632, 1980). Based on this result, ORF 4 encodes the  $\alpha$ -subunit for riboflavin synthase, which catalyzes the final step of the biosynthetic pathway: dismutation of 6,7-dimethyl-8-ribityllumazine to riboflavin (Figure 1, structures 5 and 6, respectively) and 5-amino-6-ribitylamino-2,4-(1H, 3H)-pyrimidinedione.

10. The remaining enzymatic steps in riboflavin synthesis were tentatively assigned to coding regions by aligning the position of ORFs to a physical map of rib mutations in the operon (Morozov et al., Mol. Genet. Mik. Virusol. no. 7:42 (1984)). Mutations for defective GTP cyclohydrolase were reported to map to the 0.5 kb HindIII fragment. Since ORF 3 encompasses this restriction fragment, we concluded that ORF 3, at least in part, encodes this enzymatic function, which catalyzes the formation of 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone-5'-phosphate from GTP (Figure 1, structures 2 and 1, respectively). In addition, the biosynthetic gene encoding a rib-specific reductase was reported to be contained entirely within the approximately 1.8 kb HindIII fragment. Since this fragment contains only two complete coding regions, the  $\beta$  subunit of the riboflavin synthase gene and ORF 2, we speculate that ORF 2 encodes the reductase, which catalyzes the formation of 5-amino-6-(ribitylamino)-2,4(1H,3H)-pyrimidinedione-5'-phosphate from 5-amino-6-(ribosylamino)-2,4(1H,3H)-pyrimidinedione-5'-phosphate from 5-amino-6-(ribosyla

In addition, a similar rho-independent transcription termination site was detected in the apparent leader region of the operon, downstream of the putative  $\sigma^A$  P<sub>1</sub> promoter but just upstream of the first coding region of the operon, ORF 5 (Figures 4 and 9). This potential terminator structure may be involved in regulation of the rib operon by a termination/ anti-termination mechanism. In addition, a roseoflavin-resistant (R<sub>o</sub>F<sup>R</sup>) dependent regulatory region is present on a 0.7 kb Sall-BgIII restriction fragment of ORF3.

## Assignment of rib ORFs to Protein Products

One method for confirming whether the rib-specific ORFs encode proteins is to "visualize" the size and number of proteins synthesized from the cloned DNA in an S-30 in vitro coupled transcription/translation reaction using PRF2 and its various derivatives as templates. The S-30 fraction kit (New England Nuclear; used according to manufacturer's specifications) is especially efficient in translating B. subtilis genes due to the presence of their strong ribosome binding sites.

Using the cloned 10 kb EcoRl fragment of pRF2 or pRF4 as templates, we expected to detect five putative rib-specific proteins: \$\beta\$ riboflavin synthase, 14.7 kilodaltons (kd) (Ludwig et al., J. Biol. Chem. 262:1016, 1987); and the proteins from ORF 2, 13.6 kd; ORF 3, 43.7 kd; ORF 4, 23 kd; and ORF 5, 39.7 kd. We also expected to detect at least two other proteins, encoded by ORF 6 (11.6 kd) and ORF 1 (at least 18.7 kd), as well as any additional proteins encoded by genes present in the unsequenced regions of the 10 kb cloned DNA fragment. In addition, vector-associated proteins, including the bla and cat antibiotic resistance gene products, were also expected (the tet gene is not strongly expressed in S-30 reactions).

Excluding the bla- and cat-specific proteins (32 kd and 18 kd, respectively) and other vector-associated proteins, a total of six major <sup>35</sup>S-labelled proteins were detected, with molecular weights of 47 kd, 44 kd, 38 kd, 26 kd, 20 kd and 15 kd, on a 15%-SDS polyacrylamide gel of the S-30 reactions with pRF2 or pRF4. To assign these protein products to their corresponding rib-specific ORFs, S-30 reactions were repeated using various available deletion derivatives, cat-insertion derivatives, and subcloned fragments of the 10 kb EcoRl cloned DNA (Figure 10). The results are shown in Table V.

50

40

5

10

15

25

#### TABLE V

Г	RIB-SPECIFIC PROTEINS OBSERVED IN S-30 REACTIONS							
L	1110 01 1	01110111012						
	Plasmid	47,000	44,000	26,000	15,000			
-		Daltons (ORF	Daltons (ORF	Daltons (ORF	Daltons (ORF			
		3)	5)	4)	2)			
Γ	pRF2	+	+	+	+			
-	pRF4	+	+ -	+	+			
I	pRF21	-	-	+	-			
İ	pRF5	-	-	-	+			
·	pRF29	-	-	•	-			
	pRF12	+	-	+	+			
	pRF10	-	-	-	-			
-	pRF38	-	-	-	-			
	pRF24/pRF20	-	+	· <b>+</b>	+			
İ	pRF23	+	-	+	+			

Based on these results, protein products were assigned to ORF 3 (47 kd); ORF 5 (44 kd); ORF 4 (26 kd); and ORF 2 (15 kd), with molecular weights in close agreement with the predicted sizes.

The assignment of products to ORF 2 and the  $\beta$  riboflavin synthase gene were less straightforward than the assignments to the other ORFs. Since the S-30 reaction of pRF2 produced a 15 kd protein which was close to the predicted size of the proteins encoded by either gene, it was first assumed that this protein band actually contained both protein species. However, the cat insertion into ORF 2 in plasmid pRF38 completely removed this protein band, replacing it with a much smaller protein of 6 kd, which is in close agreement with the predicted size of the truncated ORF 2. Based on these results, the 15 kd protein appears to be generated only by ORF 2. It is not clear why the  $\beta$  riboflavin synthase protein is not visualized on the gels of the S-30 reactions. Taken in total, however, the results confirmed the existence of five rib -specific coding regions: ORF 5, ORF 4, ORF 3, ORF 2 and the  $\beta$  riboflavin synthase gene.

In addition, ORF 1 appeared to encode a 38 kd protein, while no product was identified for ORF 6.

#### Regulatory Mechanisms of the rib Operon

In B. subtilis, a recurring pattern of gene organization and regulation for biosynthetic pathways has been observed by several investigators. The nucleotide sequences of the tryptophan biosynthetic pathway (Henner et al., Gene 34:169, 1984) and the de novo purine nucleotide pathway (Ebbole and Zalkin, J. Biol. Chem. 262:8274, 1987) of B. subtilis both contain clustered, overlapping genes transcribed as a polycistronic message and regulated at least in part by a novel transcription termination/anti-termination mechanism, involving a repressor protein which can be encoded by a gene unlinked to the biosynthetic operon (Zalkin and Ebbole, J. Biol. Chem. 263:1595, 1988). Since we found that the organization of the rib biosynthetic and regulatory genes is strikingly similar to those of the B. subtilis trp and pur pathways, we hypothesized that the rib operon might be regulated, at least in part, in a similar manner.

Briefly, the key characteristics of the transcription termination/anti-termination model include (Shimotsu et al., J. Bacteriol. 166:461, 1986): (i) the presence of a long 5 leader sequence that precedes the first gene in the operon; (ii) the presence in the RNA leader of two or more overlapping dyad symmetries that have the potential to form mutually exclusive RNA stem-loops, one structure functioning as a rho-independent transcription terminator and the other as an "anti-terminator" (blocking the formation of the rho-independent transcription terminator); (iii) under repressive conditions, the repressor protein, activated by the end product of the pathway, binds to the nascent mRNA at a site which prevents formation of the anti-terminator, thus allowing formation of the terminator which terminates transcription; (iv) under derepressive conditions, binding of the unactivated repressor is precluded, resulting in the formation of the anti-terminator causing read-through transcription into the coding region of the operon.

As discussed above, the most likely site for initiation of transcription in the rib operon is a  $\sigma^A$  promoter, P<sub>1</sub>, located about 290 bp upstream from the first gene in the operon. Preliminary analysis of the RNA leader sequence indicated that it contained most, if not all, of the structures required for regulation by the

5

10

termination/anti-termination model. Within this region, a stem-loop structure followed by a string of thymidines resembling a rho-independent transcription terminator was identified approximately 50 bp upstream of ORF 5; this sequence has the potential to form a hairpin with a  $\Delta G$  of -26 kcal/mol (Figure 9). In addition, several potential stem-loop structures with  $\Delta G$ 's ranging from -13 to -16 kcal/mol were located within the rib 5 leader that could possibly qualify as the anti-terminator sequence.

In addition to the primary site for the initiation of transcription, usually located upstream from the first gene in the operon, there exist in some biosynthetic pathways secondary promoter sites located within the internal regions of the operon. The possibility of there being a second promoter site within the rib locus was also suggested by previous R-loop heteroduplex studies of the rib operon (Osina et al., FEBS Letters 196 :75-78, 1986), showing two or more sites for the initiation of mRNA synthesis. Our preliminary analysis of the intercistronic gaps of the rib operon did not detect such secondary promoter sites. However, when this analysis was extended to all of the sequences within the operon, another  $\sigma^{A}$  promoter,  $P_{2}$ , was identified within the 3 end of ORF 4, just downstream from the Sall restriction site (Figure 4). Thus it is possible that the expression of ORF 2, ORF 3, and the  $\beta$ -subunit for riboflavin synthase is also under the control of this secondary promoter. In addition, a possible third  $\sigma^{A}$  promoter,  $P_{3}$ , was identified just upstream of ORF 2. Therefore ORF 2 is possibly also under the control of this additional promoter.

The location of putative coding regions, promoters and transcription termination sites on the DNA sequence of the 5.5 kb B. subtilis rib-specific region is shown in Table VI.

20

25

30

35

40

45

50

TABLE VI

CODING REGIONS, PROMOTER, AND TRANSCRIPTION TERMINATION SITES OF THE B. SUBTILIS RIB OPERON							
		bp Number <sup>a</sup>					
Coding Regions	ORF 6 ORF 5 ORF 4 ORF 3 \$\beta\$ riboflavinsynthase gene ORF 2 ORF 1	364-678 1101-2183 2197-2841 2859-4052 4088-4549 4665-5036 5567-5057 <sup>b</sup>					
σ <sup>A</sup> Promoters	P <sub>1</sub> P <sub>2</sub> P <sub>3</sub>	771-799 2528-2556. 4545-4574					
rho -Independent Termination Sites	Upstream from 5' promoter  Within 5' leader RNA  At 3' end of <u>rib</u> operon	708-748 1034-1067 5038-5090					

<sup>&</sup>lt;sup>a</sup> of Figure 3.

Example 8: Construction of vectors containing a modified rib operon

The above functional analysis of the rib operon of Bacillus subtilis for the first time delimiting the regulatory regions and open reading frames in the nucleotide sequence permits construction of new vectors which are useful for increasing the yield of riboflavin production. The knowledge of the location of the specific genes required for riboflavin biosynthesis, of the location of transcriptional control regions, and other relevant regions (e.g., RBS) in those genes allows changes in such regions to be made. There follow a few examples of such manipulations.

Construction of an integration plasmid with a rib operon on a smaller DNA fragment

<sup>&</sup>lt;sup>b</sup> Coding region oriented in opposite direction.

The integrating vector used to construct the riboflavin overproducing strain RB50::[pRF8] contains a 10 kb EcoRl fragment including the rib operon. Since the rib operon appears to occupy less than 6 kb of DNA a new integration vector was constructed (pRF40) containing the rib operon on a smaller DNA fragment. The smaller size of this clone allows higher amplification of rib genes resulting in higher yields of riboflavin.

Referring to Fig. 12, pRF40 was constructed from pRF36 which is a plasmid in which the 0.8 kb Ncol fragment of pRF2 is replaced with a cat gene. The rib operon is contained on a 6.5 kb Xbal-EcoRl fragment. This fragment was isolated and ligated to pUC19 (Yanisch-Perron et al., Gene 33, 103, 1985; available from New England Biolabs, Boston, MA, USA and Bethesda Research Laboratories, Maryland, USA) digested with Xbal and EcoRl. The ligated DNA was transformed into DH5 $\alpha$  E. coli and plated onto LB plates containing 40  $\mu$ g/ml X-gal and 50  $\mu$ g/ml ampicillin. Analysis of miniprep DNA prepared from white colonies indicated that pRF39 contained the 6.5 kb Xbal-EcoRl fragment.

pRF39 was digested with EcoRI, treated with CIAP, and then ligated to a 1.6 kb EcoRI fragment containing the cat gene. The ligated DNA was then transformed into DH5 $\alpha$  E. coli and appropriate colonies selected for plating onto LB + 10  $\mu$ g/ml chloramphenicol; two colonies were chloramphenicol-resistant. Analysis of miniprep DNA prepared from these colonies confirmed the presence of the cat gene. One of these plasmids is pRF40 (Fig. 14).

## Construction of plasmids containing transcriptionally modified rib operon

As described above, it is useful to replace the promoter and operator regions of the riboflavin operon with promoters allowing constitutive expression of the riboflavin biosynthetic genes. Plasmids containing such constructs can then be used to produce bacterial strains which will produce increased levels of riboflavin. A few examples, not limiting in the invention, are provided below.

Referring to Fig. 13, the riboflavin promoter and regulatory region were removed and replaced with an SPO1 promoter. We took advantage of the BgIII site located at position 1130 at the start of ORF3. Oligonucleotides were synthesized (RB5 and RB6, see Fig. 18) that recreated the DNA sequence 5 to the BgIII site (the first few amino acids of ORF5 and the SD sequence) up to position 1058. Reconstruction of the 5 end of the operon stopped before any of the proposed DNA regulatory structures (Fig. 13). At their 5 ends the oligonucleotides contained BamHI, NsiI, and EcoRI restriction sites, allowing for placement of various promoters 5 to the rib operon. Because of the various restrictions sites in the rib operon it was necessary to construct the operon with the new promoters in several steps, as follows.

A 1.4 kb Sall-BgIII fragment was isolated from pRF36 (Fig. 13). This fragment was ligated with the two oligonucleotides and EcoRI-Sall-digested pUC19. The ligated mixture was then transformed into E. coli DH5 $\alpha$  cells and plated onto LB containing 50  $\mu$ g/ml ampicillin and 40  $\mu$ g/ml X-gal. Minipreps were prepared from Apr white colonies; one plasmid having the desired structure is pRF46 (Fig. 13).

pRF46 was digested with BamHl and Sall and the 1.4 kb fragment isolated. This fragment was then ligated with the 400 bp EcoRl-BamHl fragment of pNH202 (pUC8 containing the SPO1-15 promoter, Lee and Pero, J. Mol. Biol., 152 :247-265, 1981) and pUC19 cut with Sall and EcoRl. The ligated DNA was then transformed in DH5 $\alpha$  E. Coli, which were plated onto LB + ampicillin + X-gal. Miniprep DNA was prepared from white colonies; and pPRF48 had the desired structure (Fig. 13).

pRF48 was digested with EcoRI and Sall and the 1.8 kb fragment isolated. This fragment was ligated with the 4.0 kb Xbal-Sall fragment (containing the rest of the rib operon) from pRF2 and Xbal, EcoRI-cut pUC19. The ligated mixture was then transformed into E. coli DH5 $\alpha$  cells which were plated on LB + ampicillin + X-gal. Miniprep DNA was prepared from white colonies; pRF49 had the desired structure, and supernatants from culture containing this plasmid was yellow, indicating riboflavin production (Fig. 13).

To place the cat gene in pRF49, to allow selection in B. subtilis, the plasmid was digested with Xbal and ligated to a 1.3 kb cat-containing Xbal fragment from pEcc4. The ligated DNA was transformed in E. coli DH5 cells. Hundreds of Ap $^{\rm r}$  colonies resulted, and the colonies were patched onto plates containing LB + 10  $\mu$ g/ml chloromphenicol. Approximately 10% of the colonies grew on the chloramphenicol plates, indicating the presence of the cat gene. One cat-containing plasmid is called pRF50 (Fig. 14).

The above example shows placement of a new promoter upstream of ORFs. We found that it is also useful to place a promoter after  $P_2$  between ORF3 and ORF4 in order to further increase riboflavin production. An example of such construction now follows.

Referring to Figs. 14 and 15, to place a copy of the SPO1-15 promoter upstream of ORF3 we made use of the restriction sites adjacent to the ORF4-ORF3 junction. The Clal site at position 2767 is located at the end of ORF4 and is unique in the rib operon. Another useful restriction site near the beginning of ORF3 is the Dral site at position 2892. Oligonucleotides were synthesized that recreated the sequence from the

above-mentioned Dral site past the start of ORF3 and placed a unique BaMHI site before the beginning of ORF3 (linkers P2-A and P2-B, Fig. 18). Another set of oligonucleotides recreated the sequence from the Clal site past the end of ORF4 and placed an ECORI site at that location (linkers P2-CII and P2-DII, Fig. 18). The SPO1-15 promoter, located on a EcoRI-BamHI fragment, was then be placed between the BamHI and EcoRI sites created by the oligonucleotides. The entire operon was put together with this additional SPO1-15 promoter as follows.

Referring to Fig. 15, the 750 bp Sall-BgIII fragment containing the ORF4-ORF3 function was subcloned to pIC2OR (Marsh et al., Gene 32, 481-485, 1984). The resulting plasmid, pRF57, was then digested with Dral and BgIII, and the predicted 270 bp Dral-BgIII fragment was isolated. This fragment and linkers P2-A and P2-B were ligated to pIC2OR cut with Sall and BgIII. The linkers placed BamHI and Sall sites upstream of the 5 end of ORF3. (The sall site was chosen for convenience since BgIII and BamHI sites are compatible and will be removed later.) The ligation was transformed into E. coli DH5 $\alpha$  cells. Plating onto LB medium + Amp and X-gal resulted in white colonies; pRF58 had the desired structure. The 330 bp BgIII-Sall fragment from pRF58 was isolated and ligated with 3.3 kb BgIII-Xbal fragment containing the 3 end of the rib operon from pRF36 (Fig. 12) and pUC19 cut with Xbal and Sall. The ligated DNA was then transformed into E. coli DH5 $\alpha$  cells, resulting in white colonies; pRF62 (Fig. 15) had the desired structure. For convenience, the 3.6 kb BamHI-Xbal fragment was isolated from pRF62 and subcloned into BamHI-Xbal-cut pUC19 (pRF64, Fig. 15). This plasmid now contained the 3.6 kb 3 end of the rib operon with an engineered BamHI site preceding ORF3.

To place the SPO1-15 promoter in front of the 3'-half of the rib operon containing the last three open reading frames, we digested pRF64 with EcoRI and BamHI and ligated it to a 400 bp EcoRI-BamHI fragment containing the SPO1-15 promoter. The ligated DNA was transformed into E. coli DH5 cells and miniprep DNA was prepared; pRF65 has the desired structure.

The SPO1-15 promoter was than engineered to place a Clal site upstream of the promoter to reconstruct the end of ORF4. The EcoRl-BamHI fragment from pNH202 containing the SPO1-15 promoter was ligated with linkers P2-ClI and P2-DII and pCI2OR-digested with BamHI and Clal. The ligated DNA was then transformed into E. coli DH5 $\alpha$  cells. White colonies resulted and miniprep analyses indicated that pRF63 had the desired structure. The 470 bp Clal-BamHI fragment was isolated then from pRF63 and ligated to the 2kb EcoRl-Clal fragment from pRF49 containing the SPOI-15 promoter and the 5<sup>-</sup>-end of the rib operon and pRF64 (Fig. 15), containing the SPO1 promoter and the 3<sup>-</sup>-end of the operon, digested with EcoRI and BamHI. The ligated DNA was then transformed into E. coli DH5 $\alpha$  cells. Miniprep DNA was prepared; pRF66 had the desired structure. In addition, E. coli containing pRF66 produced small amounts of riboflavin on LB medium + ampicillin plates, confirming that the operon was still intact.

The last step was to ligate the cat gene into the unique Xbal sites of pRF66 as described above. The resulting plasmid, pRF69 (Fig. 15) contained the cat gene in the same direction as the rib operon.

To construct a plasmid containing the entire operon with the natural or wild-type ribP<sub>1</sub> promoter and the SPO1-15 promoter after ribP<sub>2</sub>, the 6.3 kb EcoRl-BamHI fragment of pRF64, the 2.75 kb EcoRl-Clal fragment of pRF36, and the 470 bp Clal-BamHI fragment of pRF63 were ligated and tranformed into E. coli DH5 $\alpha$  cells. About 50% of the Apr colonies were yellow, indicating ribflavin production. Miniprep DNA was prepared from yellow colonies and pRF68 had the desired structure (Fig. 16). A cat gene was added to pRF68 at the Xbal site, as discussed above, to generate pRF71 (Fig. 16). This plasmid contained the cat gene in the same direction as the rib operon.

As another example of the construction of useful plasmids in this invention, there now follows an example in which one or more promoters can be introduced within the riboflavin operon without prior removal of existing DNA sequences.

As an example, a prototype modified operon was constructed in pRF78, which contains a single copy of the SPO1-15 promoter inserted within a 30 bp non-essential region located between ribP<sub>1</sub> and a putative rho-independent transcriptional termination site (Fig. 14), an inactivated ribP<sub>1</sub> promoter to prevent possible transcriptional interference of the SPO1-15 promoter, an active ribP<sub>2</sub> promoter, the five structural genes encoding rib biosynthetic enzymes, and approximately 1.5 kb of flanking DNA nucleotide sequences downstream from the end of the riboflavin operon.

Referring to Fig. 14, the 1.7 kb Ncol-Pstl fragment of pRF2, a fragment that contains the 5 promoter region of the rib operon and flanking regions, was first subcloned into mp19, a derivative of the E. coli bacteriophage vector M13 (United States Biochemical Catalog, 60-61, 1987; available from New England Biolabs, Massachusetts, USA). One recombinant phage, M1.7, was recovered and standard DNA sequence analysis of the promoter region revealed a spontaneous mutation of the -10 region of the ribP<sub>1</sub> promoter, a TA-to-CT change, which may inactivate the promoter. Single stranded DNA was prepared and annealed to a synthetically-generated 55 bp DNA oligomer (see Fig. 17), containing a combination of restriction enzymes

sites, 5'-EcoRl-Smal-BamHl-3', flanked on either side by additional sequences homologous to the DNA region upstream from ribP<sub>1</sub>. Double-stranded DNA molecules were synthesized using standard site-directed mutagenesis (SDM) protocols. These DNA molecules were introduced into the E. coli host TG-1 (Amersham Corp., Illinois, USA) by transfection to generate recombinant phage plaques. One recombinant phage was found to contain the desired modified DNA sequence, as determined by standard DNA sequence analysis.

The modified rib promoter region was then rejoined to the rib structural genes of the operon using a pair of unique Nsil restriction enzymes sites 750 bp apart of the flank the ribP<sub>1</sub> region and surrounding sequences. Double-stranded DNA molecules of the phage recombinant were prepared, digested with Nsil, the 750 bp fragment isolated, and the fragment ligated to dephosphorylated, 8.7 kb Nsil fragment of pRF39 $\Delta$ R1 (a plasmid derived from pRF39, Fig. 12, that contains the wild-type rib operon). The ligated DNA molecules were introduced into E. coli DH5 $\alpha$  cells by transformation, selecting for ampicillin-resistance, which resulted in the recovery of an Apr colony harboring the desired recombinant plasmid, pRF75.

The SPO1-15 promoter was next inserted upstream from ribP<sub>1</sub> by digesting pRF75 with a combination of EcoRI and BamHI enzymes, ligating the cut DNA to purified 400 bp EcoRI-BamHI SPO1-15-containing restriction fragment, and introducing the ligated DNA into E. coli DH5α cells by transformation, selecting for ampicillin-resistance. One Apr colony was found to harbor the recombinant plasmid, pRF77, containing the desired SPO1-15-modified rib operon. A chloramphenicol-resistance gene, cat, on a 1.6 kb Xbal restriction fragment, was subsequently introduced into pRF77 at the unique Xbal site, generating plasmid pRF78 (Fig. 14).

This prototype operon was further modified to contain an active ribP<sub>1</sub> promoter, and/or a second copy of the SPO1-15 promoter introduced downstream from ribP<sub>2</sub> within a intercistronic region between the rib coding regions ORF3 and ORF4, as described above. For example, plasmid pRF88, containing a derivative of the modified rib operon in pRF78 with an active ribP<sub>1</sub> promoter (Fig. 14) was constructed by the same procedure described above, using a recombinant phage containing the wild-type ribP<sub>1</sub> promoter. In other examples, a second copy of the SPO1-15 promoter, located downstream from ribP<sub>2</sub>, was inserted into the existing modified rib operon-containing plasmids pRF78 and pRF88 by removing the 2.0 kb Bglll fragment of either plasmid DNA and inserting the 2.4 kb Bglll fragment of pRF66, generating plasmids pRF81 and pRF89 respectively (Fig. 14).

### Construction of Ade RB50 strains

It is important to use strains of bacteria that require as few components to be added to a fermentation medium as possible. Such strains are cheaper to ferment in order to produce riboflavin. To this end, adenine revertants which contained amplified modified rib operons were constructed. These revertants may not be true revertants of pur-60, but rather include mutations at another site which suppresses the requirement for adenine. As discussed below they produce about 25% more riboflavin than the non-reverted strains. Examples of such constructions are now described.

Plasmids pRF8, pRF40, pRF50, pRF69, pRF71, pRF78, pRF81, pRF88 and pRF89 were each transformed into RB50 (a RoF′, deregulated B. subtilis strain) selecting for chloramphonicol resistance (Cm′). A resistant colony was chosen for each strain. Ade \* revertants of each strain was isolated by growing bacteria in RMM1 broth containing 10 μg/ml adenosine, and plating samples of the cultures onto minimal agar plates. One colony from each Ade \* strain was selected and the vector DNA was amplified by selecting colonies that grow on increasingly higher levels of chloramphenicol, to a maximum level of 60 μg/ml.

#### Second site Integration

As described above, it is important to amplify an engineered rib operon in the B. subtilis chromosome to achieve high titers of riboflavin. It is also important to ensure that the number of DNA copies of the rib operon within a chromosome are not limiting to riboflavin production. Further amplification of the rib operon can be achieved by integrating and amplifying copies of the rib operon at more than one site in the B. subtilis chromosome to further increase riboflavin yield. One example of how such second site integration can be achieved is described below.

The above described vectors have all relied upon the cat gene to allow integration at the site of the rib operon. In order to insert the rib genes at a second site, it is preferable to have a different antibiotic resistance gene for use at that second site. For example, a tetracycline-resistance (tet) from B. subtilis can be used (Perkins and Youngman, J. Bacteriol., 155:607-615, 1983). Such tet genes are well known to those

30

of ordinary skill in the art and are readily available to such persons. In one such construction, for example, pRF78 (Fig. 14), which contains a modified version of the rib operon, the plasmid can be cut with Xbal and ligated to a 2.4 Xbal fragment containing the tet gene. The resulting plasmid contains the tet gene at the Xbal site and is called pRF85.

A strain which is deleted for the entire rib operon and which has a tet gene integrated at a second site is required to cause integration of pRF85 at that site. One such site is the bpr gene encoding bacillopeptidase F, a minor non-essentiall extracellular protease. An E. coli plasmid containing the bpr gene, pKT2, (Sloma et al., J. Bacteriol., 172:1470-1477, 1990) was digested with EcoRV. This EcoRV site is in the coding region of bpr. The DNA was then ligated to a 2.4 kb Eco RI fragment containing the tet gene that had been blunt-ended. The resulting plasmid (containing the tet gene at the EcoRV site of bpr) was called pKT2-tet. This DNA was linearized with EcoRI and then transformed into RB52, a strain deregulated for riboflavin synthesis. Tet colonies resulted and one such colony was called RB54. The integrated tet gene at bpr will function as a homologous sequence for the integration of pRF85.

To ensure that the cloned riboflavin operon of pRF85 would be inserted at the second chromosomal site containing the tetracycline-resistance gene, a region containing the original riboflavin operon and flanking DNA, equalling that contained in pRF85, was deleted from the chromosome of RB54 by in vitro methods. Briefly, this involves first generating an E. coli recombinant plasmid where the cloned riboflavin operon and flanking regions between the Ncol and Xbal restriction sites are removed and replaced by a chloramphenicol-resistance gene, cat, that is expressed in B. subtilis bacteria. This plasmid is then used to delete the chromosomal riboflavin operon by transforming RB54 with linearized plasmid molecules and selecting for chloramphenicol resistant (Cmr) bacteria. Cmr bacteria result from a recombinant event (marker-replacement) which replaces the wild-type rib genes with the deleted copy containing the cat gene.

Specifically, plasmid pRF34 (see example 6) was used to generate an E. coli plasmid containing an in vitro-generated riboflavin operon deletion. This plasmid is derived from pRF2 where the riboflavin operon is flanked on either end by two unique Xbal sites (one site located upstream from the 5'-end of the rib operon next to the deleted 0.8 kb Ncol fragment and the second site located approximately 1.6 kb downstream from the end of the operon) and a cat gene is inserted outside of this region. By digesting pRF34 with Xbal and ligating the cut DNA molecules under dilute DNA concentrations, a recombinant plasmid, pRF82, was recovered where a 7.2 kb region containing the riboflavin operon is removed and essentially replaced with the cat gene. Plasmid pRF82 was linearized by restriction enzyme digestion and the cut DNA used to remove the chromosomal riboflavin operon of RB54 by DNA transformation, selecting for Cmr bacteria, resulting in marker replacement. Cmr colonies were screened for riboflavin auxotrophy and one Rib-Cmr colony, RB55, was recovered for further investigation.

Plasmid pRF85 was transformed into strain RB55, selecting for Rib\*. One Rib\* transformant was chosen and called RB58. This strain has the rib operon integrated at bpr by homologous recombination between the tet<sup>r</sup> genes in the plasmid and the chromosome. Chromosomal DNA from RB58 was prepared and can be used to transform RB50::[pRF69], selecting for Tet'. These resistant colonies will then have the modified rib operon integrated at the site of the rib operon and at bpr. The rib operon integrated at rib will be amplified by selecting for colonies that grow in the presence of increasing levels of chloramphenicol as described above, and the second copy of the rib operon will be amplified by selecting colonies that grow on increasing levels of tetracycline.

#### Example 9: Fermentative Production of Riboflavin

Evaluation of riboflavin-overproducing strains was conducted in Chemap 14-liter vessels in carbon-limited fed-batch fermentations, with riboflavin content measured by HPLC. Since enzymes encoded by the genes for riboflavin synthesis are rate-limiting, the rib genes, which were amplified, were maintained at high-copy number by the inclusion of  $60~\mu g/ml$  chloramphenicol in the inoculum seed train, but not in the fermentor.

A culture of B. subtilis RB50::[pRF69] was grown on Tryptose Blood Agar Base (TBAB Difco) containing 60  $\mu$ g/ml of chloramphenicol (CAM). Colonies were transferred to 300 ml baffled flasks containing 25 ml of riboflavin minimal medium (RMM; containing sodium glutamate 2.0 g/l, Casamino acids (Difco) 0.2 g/l, Yeast extract (Difco) 0.2 g/l, KH<sub>2</sub>PO<sub>4</sub> 6.0 g/l, K<sub>2</sub>HPO<sub>4</sub> 14.0 g/l, (NH<sub>4</sub>)  $_2$ SO<sub>4</sub> 2.0 g/l, sodium citrate 1.0 g/l, MgSO<sub>4</sub>•7H<sub>2</sub>O 0.2 g/l, adenosine 0.05 g/l, glucose 15.0 g/l, pH 7.0) with 60  $\mu$ g/ml CAM. The inoculated flasks were incubated by shaking at 250 rpm and 37 °C. After 8 hours, sterile glycerol was added to a final concentration of 15% and 1 ml aliquots were stored at -80 °C.

In order to initiate a fermentation a frozen vial of RB50::[pRF69] was thawed at 37°C and transferred

45

into a 300 ml baffled flask with 25 ml of RMM with  $60~\mu g/ml$  CAM and shaken at 250 rpm and  $37~^{\circ}$  C. After 8 hours, 6 ml of the growing culture was used to inoculate 300 ml of fermentation medium (see Table VII below) in a 2 liter transfer flask. Such flask contained 300 ml of fermentation medium to which had been added 90 ml of a mixture of 15% glucose and 30% maltose. Chloramphemicol was added to a final concentration of  $60~\mu g/ml$ . After incubation for 12 hours at 200 rpm on a shaker with a 2" diameter orbit at  $37~^{\circ}$  C, the contents of such a flask was transferred to 7 liters of fermentation medium in a 14 liter fermentation vessel.

During fermentation, the broth was continually monitored for pH and dissolved oxygen (DO<sub>2</sub>). Off gas was continuously analyzed by quadrapole mass spectrometry and carbon dioxide evolution (CER) and oxygen uptake rates were recorded.

A comparison of several fermentations demonstrated the reproducibility of the control systems. The initial carbohydrate was exhausted from fermentation with RB50::[pRF8] $_{50}$  after 4 hours of growth, causing a rise in pH and a fall in CER. At that point, carbohydrate feeding was initiated and logarithmic growth resumed until DO $_2$  became limiting at 6 hours. The rate of carbohydrate feeding was computer-controlled to maintain the DO $_2$  between 10-20% of saturation throughout the remaining fermentation time.

Excess carbohydrate in the fermentors does lead to oxygen starvation and reduced riboflavin production. Oxygen transfer limitations determine the duration of logarithmic growth, final cell density and the riboflavin production rate. To increase the oxygen transfer rate, Chemap fermentors were run at 1000 rpm with a head pressure of 0.6 atmospheres.

Supplementation of the medium carbohydrate feed with yeast extract led to an increase in riboflavin production as compared to media without supplementation (Figure 11, open squares: RBF-14; Table VII). However, because of its high cost, the amount of yeast extract was systematically reduced by substituting less expensive, inorganic ingredients. Substitution of ammonium hydroxide for sodium hydroxide in pH control allowed a reduction of yeast extract in the feed and resulted in an increase in both cell mass and riboflavin titer (Figure 11, closed squares: RBF-22; Table VII). Fermentation times were also reduced. In other fermentations, moreover, yeast extract was completely eliminated from the feed and replaced with a combination of inorganic salts of ammonium and phosphate, resulting in a further increase in riboflavin production and a reduction of process time (Figure 11, open circles: RBF-23; Table VII).

The original RB50::[pRF8]<sub>60</sub> was auxotrophic for adenine because of its pur-60 mutation. When experiments were conducted to determine the minimum amount of adenosine required by the strain, in order to minimize its inhibition of earlier biosynthetic enzymes involved in the pathway leading to the riboflavin-precursor IMP (Figure 2), RB50::[pRF8]<sub>60</sub> (and, in general, RB50 strains with a rib operon amplified within their chromosome) was found to be unstable in its adenosine requirement and prototrophic revertants (Ade<sup>\*</sup>) were produced at a fairly high frequency. In shake flasks, the Ade<sup>\*</sup> revertants appeared to grow and produce riboflavin at least as well as the RB50::[pRF8]<sub>60</sub> parent. When evaluated in fermentors, the revertant, RB50::[pRF8]<sub>60</sub>(Ade<sup>\*</sup>), did not require adenosine in the media formulation. More importantly, the revertant grew at a faster rate and produced 25% more riboflavin than its parent strain in less time. A titer of 5.4 g/l riboflavin was produced in 49 hours (Figure 11, closed circles: RBF-29; Table VII). In additional fermentations, moreover, Hy Soy T was removed from the initial charge or medium and replaced with corn steep liquor resulting in a further increase in riboflavin production to 6.3 g/l in 48 hours. (RBF-42, Table VII).

Under these fermentation conditions, further significant increases in riboflavin production were demonstrated using bacterial strains that contained engineered riboflavin operon DNA. Strains containing the wild-type riboflavin operon on a 6.5 kb EcoRl-Xbal restriction fragment, RB50::[pRF40]60(Ade<sup>+</sup>), produced 7.4 g/l of riboflavin in 48 hours. Moreover, strains containing a transcriptionally-modified rib operon where the ribP1 promoter and regulatory region were replaced by the constitutive SPO1-15 promoter, RB50::[pRF50]60(Ade<sup>+</sup>), produced 9.0 g/l of riboflavin in 48 hours. These results demonstrate that modification of the riboflavin operon through the removal of regulatory regions and/or through the introduction of stronger, constitutive exogenous promoters leads to increases in riboflavin titer.

50

#### **TABLE VII**

	FERME	NTATION C	OMPONENT	S AND COND	ITIONS	
	Component	RBF-14	RBF-22	RBF-23	RBF-29	RBF-42
	Initial Charge (g/l)					
	Glucose	10.00	15.00	15.00	15.00	15.00
	Corn step liquor					10.00
1	Hy Soy T	15.00	15.00	15.00	10.00	
	Sodium glutamate				5.00	5.00
	Amberex 500	15.00	15.00	20.00	20.00	20.00
	KH₂PO₄	5.00	5.00	7.50	7.50	7.50
	MgGl <sub>2</sub> .6H <sub>2</sub> O	0.5	0.5	1.50	1.50	1.50
	MnSO <sub>4</sub>	0.05	0.05	0.05	0.05	0.05
	Adenosine	0.05	0.05	0.05		
	MAZU DF37C	2.50	2.50	2.50	2.50	2.50
	FeCl <sub>3</sub>			0.025	0.02	0.02
	CaCl <sub>2</sub>			0.50	0.50	0.50
	ZnSO <sub>4</sub>			0.0005		
	CuCl <sub>2</sub>			0.0013		
	CoCl <sub>2</sub>		<del></del>	0.0013		
	Nutrient Feed (g/l)					
;	Amberex 500	160.00	120.00			
	NH₄Cl			7.50	7.50	7.50
	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>			7.50	7.50	7.50
	KH <sub>2</sub> PO <sub>4</sub>			15.00	15.00	15.50
	MgSO <sub>4</sub> .7H <sub>2</sub> O			2.50	2.50	2.50
	DL-70 syrup (as DS)	600.00	600.00	600.00	600.00	600.00
	pH Control Range					
	6.6	H₂SO <sub>4</sub>	H₂SO₄	H₂SO₄	H₂SO₄	H <sub>2</sub> SO <sub>4</sub>
	6.5	NaOH	NH₄OH	NH₄OH	NH₄OH	NH₃
	Conditions	ÿ				
	Air (vvm)	1.0	1.5	1.5-2.0	1.5	1.50
	RPM	1000.0	1000.0	1000.0	1000.0	1000.0
	Temp. (°C)	37.0	37.0	37.0	37.0	37.0
'	Pressure (bar)	0.5	0.5	0.5-0.75	0.6	0.6
	Riboflavin (g/l)	3.4	4.1	4.3	5.4	6.3
		(64 hrs)	(56 hrs)	(53 hrs)	(49 hrs)	(48 hrs.)
	Dry Weight (g/l)	33.6	· 36.0	36.8	ND	44.6

The kinetics of riboflavin production in the various fermentations were analyzed using the Luedeking-Piret model. In all cases, the specific productivity declined from the conclusion of the exponential growth phase to the end of fermentation. Also, it was clear that riboflavin production was growth-associated under the fermentation conditions used.

We have discovered that the yield of riboflavin can be increased by changing the fermentation components and conditions. The yield of riboflavin can be increased compared to those conditions described above using those fermentation components and conditions shown in Table VIII.

#### **TABLE VIII**

	(g/liter)
Initial Batch	
Yeast Extract Glucose KH <sub>2</sub> PO <sub>4</sub> MgCl <sub>2</sub> .H <sub>2</sub> O CaCl <sub>2</sub> .2H <sub>2</sub> O MnSO <sub>4</sub> FeCl <sub>3</sub> .6H <sub>2</sub> O Mazu DF37C	20 25 7.5 1.5 1.0 0.05 0.025 2.5
Corn Steep Liquor Sodium Glutamate	10 5
Feed Medium (3 liter used)	s total
Glucose NaCitrate KH <sub>2</sub> PO <sub>4</sub> Succinic Acid MgSO <sub>4</sub> .7H <sub>2</sub> O	583.3 6.67 15 1.67 1.67

Briefly, in one such fermentation the starting material is 6.65 liters of batch medium and 0.35 liters of bacterial (RBSO::[pRF50]<sub>60</sub>Ade<sup>\*</sup>) inoculant. Oxygen levels are monitored with a Chemap polarographic dissolved oxygen electrode. Dissolved oxygen levels are maintained at 15% ± 5% by means of computer regulated addition of the feed medium. Total feed added is about 3.0 liters in 48-56 hours. Fermentation pH is maintained at 6.5± 0.1 (using 1N H<sub>2</sub>SO<sub>4</sub> and NH<sub>3</sub> gas), and fermenter pressure is maintained at 0.6 bars, and air flow at 10.5 liters/min. Under these conditions, strain RB50::[pRF50]<sub>60</sub>(Ade<sup>\*</sup>) produced 11.0 g/l riboflavin in 48 hours, which represents an improvement in production of approximately 20% compared to the previous fermentation conditions. Finally, a further increase in riboflavin production was demonstrated using the bacterial strains RB50::[pRF69]<sub>60</sub>(Ade<sup>\*</sup>) containing a transcriptionally-modified riboflavin operon containing two SPO1-15 promoters, one replacing ribP<sub>1</sub> and regulatory sequences, and a second inserted between ORF3 and ORF4. This strain produced 13.0-14.0 g/l riboflavin in 48 hours, and 15 g/l in 56 hours, demonstrating that increased transcription of the riboflavin operon using two strong exogeneous promoters increases production levels of riboflavin.

### Deposit of Microorganism

Plasmid pRF69 has been deposited with the American Type Culture Collection on June 6, 1990 under the Budapest Treaty and was assigned Accession No. ATCC 68338.

E. coli strain DH5 containing plasmid pRF50 has been deposited with the American Type Culture collection on May 30, 1990 under the Budapest Treaty and was assigned Accession No. ATCC 68332.

E. coli strain DH5 alpha containing plasmid pRF78 has been deposited with the American Type Culture collection on May 30, 1990 under the Budapest Treaty and was assigned Accession No. ATCC 68333 and Bacillus subtilis strain RB58 has been also deposited with the American Type Culture collection on May 30, 1990 under the Budapest Treaty and was assigned Accession No. ATCC 55053.

Bacillus subtilis strain RB50 was deposited on May 23, 1989 with the Agricultural Research Culture Collection (NRRL), Peoria, Illinois, under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedures, and was assigned accession number B 18502.

#### Claims

5

10

15

20

- 1. A vector comprising first a nucleic acid sequence of bacterial or yeast origin coding for one or more riboflavin biosynthetic proteins and second one or more transcription elements not naturally associated with this nucleic acid sequence.
- 2. A vector according to claim 1 wherein said nucleic acid sequence of bacterial origin is derived from Bacillus especially Bacillus subtilis or E. coli.
  - 3. A vector according to claim 1 or 2 wherein said nucleic acid sequence is coding for at least 5 riboflavin biosynthetic proteins.
  - 4. A vector according to any one of claims 1 to 3 which comprises one or two such transcription elements.
- 5. A vector according to any one of claims 1 to 4 wherein said transcription element is a promotor.
  - 6. A vector according to claim 5 wherein said promotor is a promoter chosen from a promoter naturally associated with a gene from SPO1 bacteriophage, a promoter naturally associated with veg, amy; and apr, and other transcription regulatory elements activated by the gen product of sacQ.
  - 7. A vector according to claim 6 which is pRF50, pRF69, pRF70, pRF71, pRF78, pRF81 or pRF89.
- 8. A recombinant bacterium comprising at least one copy of an exogenously introduced nucleic acid sequence at one or more sites within its chromosome, which sequence encodes one or more riboflavin biosynthetic proteins and is heritable and capable of expression by the bacterium, such that riboflavin biosynthesis by the bacterium is increased relative to a bacterium lacking such nucleic acid sequence.
  - 9. A recombinant bacterium comprising a bacterium which has been transformed by a vector according to any one of claims 1-7 whereby at least one copy of said nucleic acid sequence including said transcription elements has been introduced at one or more sites within its chromosome and said nucleic acid sequence including said transcription elements is heritable and capable of expression by the bacterium, such that riboflavin biosynthesis by the bacterium is increased relative to a bacterium lacking such nucleic acid sequence including said transcription elements.
- 10. A recombinant bacterium according to claim 9 wherein said nucleic acid sequence including said transcription elements is present in a plurality of copies at such sites.
  - 11. A recombinant bacterium according to claim 9 or 10 wherein said nucleic acid sequence including said transcription elements has been introduced at one or two sites within its chromosome.
  - 12. A recombinant bacterium according to any one of claims 8 to 11 wherein said bacterium is deregulated for riboflavin gene expression.
  - 13. A recombinant bacterium according to any one of claims 8 to 12 wherein said bacterium is an E. coli or Bacillus, especially Bacillus subtilis strain.
  - 14. A recombinant bacterium according to claim 13 wherein said B. subtilis strain is RB50 or RB58.
- 15. A process for the production of riboflavin characterized by growing a recombinant bacterium according to any one of claims 8 to 14 under suitable growth conditions.
  - 16. A process according to claim 15 wherein such suitable growth conditions are characterized by limiting the availability of a component of the growth medium in such a way that aerobic conditions for the growth of said recombinant bacterium are maintained.
- 17. A process according to claim 16, wherein said component is chosen from a carbon source, nitrogen source or a component required by said recombinant bacterium.
- 18. A process according to claim 17, wherein said component is glucose, or a carbonic acid.
- 19. A process according to claim 16, wherein said limiting step comprises limiting introduction of said component in a feed medium.
- 20. A process according to claim 19, wherein said component is glucose, or a carbonic acid.

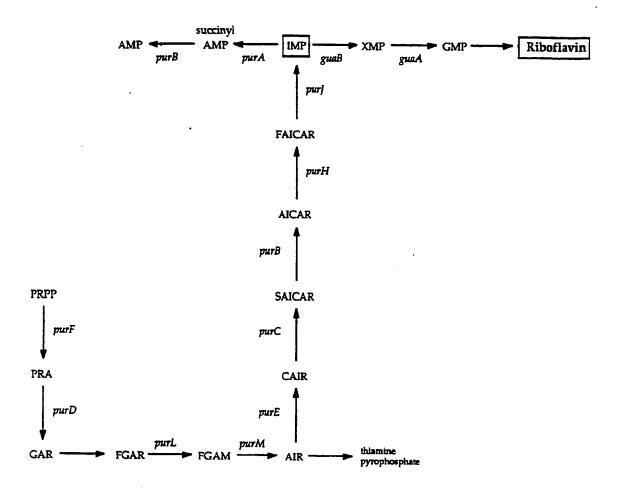
45

Fig. 1

deaminase

reductase

## FIGURE 2.



```
Page 1
      PPSANTHIXXII BXBNINDEDSXIDSBBN N 8 HSSBBNINH
SLACIANNBA (HALBPISTCSESSL A G ATESSBAA
                                 KE M
                                           SA
                                 SA N
                                            AE
      TELCHOFLAE HOMAONNAYOACAAA E L EYCAAGEE
                                 PR L
       1111211111 12141111111113 3 1 31133213
                                 21 1
                                            12
              111 111111
       / ///
                           1111 1
                                  1
     100
    MIS POLYLINER
    lqvdsrgspindsrnglglftkkqi agtsk fyfyp
     crstiedppwtavtalassrknkirvrqsifst
      egrl, rîph g q p . r p w p l h e k t n c g y v k v c f l p
     ctselpdgmstriprprkvffcfapvdfntk.
      qldvrssgghvatvakaeerffinrtr, i knev
      apra, ligupegyrgqsr, sfyfqpytltqkrg
                  BM D
                       F AB
      М
           HFH
                       M LB
                              N
                                      0
      s
           ISH
                  18 P
                            В
                              ย
                                      ĸ
                  NO N
                       ע עע
                             ν
      Ε
           NPA
                  11 1
                       H 11
           P11
     CGTTTAACGAAATGCGCTAAACAAATTAGGATCAAGCAGCTTCCCCATTGGGGCTGCTTTTTTTATATCTTTTTTACGGTCATCCCCTAAAAACAGAACAT
     GCAAATTGCTTTACGCGTTTTGTTTAATCCTAGTTCGTCGAAGGGTAACCCCCGACGAAAAAATATAGAAAAAATGCCAGTAGGGGGATTTTTGTCTTGTA
      fnemrktn.dqaashwgcffyifftvip,kqni
     rttkcakqirikqipigaaffiaflraspknrt.v.rnaqnklgsssfplgliflylfyghpikteh
     gnlsirlyf, s.aaewqpqkk, ikkvtmg, fcfm
rkvfhafcililosgmpaakkidkkrddglflv
      t.rfacflnpdllkgnpsskkyrk,p.grfvsc
                                       S AM
        н
                                       F SS
                                    SV
        ¥
                                       A EE
                                    14
                                    13
                                       N 11
     AAATTEGTATATETATAGAAAAGAAATTTTTGCAGAAATGTGAAACATATTCCCGTTATGCATCGTTATATTAATAATTTACGAGAATTTACGGTTTTTT
     300
     titaagcatatagatatettitetttaaaaacgtetttacaetttgtataagggcaatacgtagcaatataattattaaatgetettaaatgeeraaaaa
     nsy i y rke i faem. nifploivilii y entrffirisiekk flokcety sry asty... ftri y g fl
     k f v y l . k r n f c r n v k h i p v m h r y i n n l r e f t v f Y
          feyi.lfsikasihfmngnhmtinii.sfkrnk
     y î rîdîs ffnkcfhs vyer, adny. y n v li. p k k lntyry flfk q lft fcigtier, î l l k r s n v t k
```

FIGURE 3A

						Page
N	N	ХH	ĸ	s H		
S L	D	MN	N	P 6		
A	E	NF	L	0 0		
1 3	1	11	1	1 2		
			RES	ORF 6	ACATTACCATTCTCTTAC	
CATGAMAM	MGGAATAACTCATATGAA	GAATAGATTE	TATIGGETGGAGGTTAL	MATGGGAAGAATAAAAACCA		400
GTACTTTTT	TCCTTATTGAGTATACTT	ACTTATCTAAGI	TATAACCGACCTCCAAATC	TTTACCCTTCTTATTTTTGGT	TCTAATGGTAAGACAATC	
m	rnnsye gîthmn	.ids e.ih	yulev. igurfr	mgrikt k kwee.kp ngknknq	rlpfc. dyhsvs	
.sf	fsysmh	if ln hise	mnappk yqsst.	siplify fhssyfg fpfflfu	livmrn lngnq.	
		R	<b>A</b>	F	BSB M KEF M	
		S	L 	0	SES N SAO N ACA L PRK L	
		A	U	К 1	J1J 1 211 1	
		1	1	1	/ /	
			,	rtccgacagcaattggacaaac 	,	500
GAAAACGAAA L L L L f c f	ATGAACGTCCGCCAATATA  laggym ylqavic	CATGTATTTAC y i n d	TATAACTCGACTTCCTAC	MGGCTGTCGTTAACCTGTTTG  ptaigqt frqqldk;	GAACAGGAGCCTTCTTCT	500
GAAAACGAAAA lllli fcf faf sks	ATGAACGTCCGCCAATATA  laggym ylqavic tcrrly  ksapp.ii	catgtatttac y i n d : t . m v h k . i y m f h v y i	TATAACTCGACTICCTAC	AGGCTGTCGTTAACCTGTTTC	SCANCAGGAGCCTTCTTCT  ISSEE COPPK  IVIGT  VKdesss  gqgrff	500
GAMACGAAA LLLL fcf faf sks kqk akak	ATGAACGTCCGCCAATATA  laggym ylqavic tcrrly  ksapp.i .kcati vqlrny	CATGTATTTAC  y in d  t m  v h k  y m f  h v y i  t c l h	TATAACTCGACTICCTAC	ptaigqt frqqldk; sdsnwtn tgvsipcy nrccnsl esllqvf	SCANCAGGAGCCTTCTTCT  ISSEE COPPK  IVIGT  VKdesss  gqgrff	500
GAAAACGAAAA  LLLL fcffaf sks kqk bkak HME BBC	ATGAACGTCCGCCAATATA  laggym ylqavic tcrrly  ksapp.i .kcati vqlrny	catgrattrac y in d t m v h k. y m f h v y i t a l h	TATAACTCGACTICCTAC	ptaigqt frqqldk; sdsnwtn tgvaipcv nrccnsl esllqvf	SCANCAGGAGCCTTCTTCT  ISSEE COPPK  IVIGT  VKdesss  gqgrff	500
GAAAACGAAAA llll fcff f * f s k s k q k a k a k H M E B B C O O R	ATGAACGTCCGCCAATATA  laggym ylqavic tcrrly  ksapp.i .kcati vqlrny	catgrattrac y in d t m v h k. y m f h v y i t a l h	TATAACTCGACTICCTAC	ptaigqt frqqldk; sdsnwtn tgvsipcy nrccnsl esllqvf	SCANCAGGAGCCTTCTTCT  ISSEE COPPK  IVIGT  VKdesss  gqgrff	500
GAAAACGAAA  I I I I I  f c f  f a f  s k s  k q k  a k a k  H M E  B B C  O O R  2 2 2	ATGAACGTCCGCCAATATA  laggym ylqavic tcrrly  ksapp.i .kcati vqlrny  EAS N .PC A .YR E .11 3	CATGTATTTAC  y i n d  t m  v h k .  y m f  h v y i  t c l h	TATAACTCGACTICCTAC	ptaigqt frqqldk; sdsnwtn tgvaipcv nrccnsl esllqvf	ISSEEE CATCTTCAACGGAAAAA	
GAAAACGAAA  L L L L f c f i f a f i f a f i s k q k a k a k  M M E B B C O O R 2 2 2  GAATACACCAT  CTTATGTGGTA	ATGAACGTCCGCCAATATA  laggym ylqavic tcrrly  ksapp.i .kcati vqlrny  AS H .PC A .YR E .11 3 ./ TCCAGGAATATAAAGTGACG	CATGTATTTAC  y i n d  t . m  v h k .  y m f  h v y i  t c l h  CAAAATTGACGC	TATAACTEGACTICCTAC  I i e l k d v i l s . r m y . a e 9 c  S i s i f s i n l q l i y q a s p h  D	ptaigqt frqqldk; sdsnwtn tgvaipcy nrccnsl esllqvf  M B O 2	CATCTTCAACGGAAAAAA  GTAGAAGTTGCCTTTTTTT  i f n g k k s s t e k n	

FIGURE 3B

				Page :
u	H 1	4 H	BM DT M	
H S	••	S P	IB PA S	
E	-	E H	NO NO E ON	
1	3	1 1	11 11 1 TEU / STC	nslation P
TTAAATCAGGATTTATCTGATAT	AAAGAAGGTGACAAGA	TTAAGGCTTACTTCAGCAAATC	MAGCGGATCGACGGTTAAT	AAGSTTGCAAAGTGAA
	-,	,+,+	,	700
AATTTAGTCCTAAATAGACTATA	TTTTCTTCCACTGTTCT	AATTCCGAATGAAGTCGTTTAG	TTTEGEETAGETGELAATTA	HICLAACUIIIICACII
ln q d l s d i	ke g.dki	k a y f s k s	kridg. s	riqk.m
.irivli.	k k v t r	iritsanq	s g s t v n (	q g c k s e
ksgfi.y	k r r . q d	.gliqqi	K B G F F L	A V W h V 11
nf.skdsi	fspsl	ila. kild	fri	lncth
ili - ri	vfftvl	nlsveæf	. ipdvti	.pqlls
ildpniqy	lllhcs	. p k s . c i	Laserni	t a T T T
N H BM	DT	H	R HD	BM D
* *	PA	P 	S SR	IB P No N
	NO RED-INDEPENDE		A EA 1 11	11 1
1 1 11	11 TRANSCRIPTION TERMINATOR	1	1 11	P1 PROMOTER
TGATTAAAAAACATCACCTTTCG	GATCGAAGGGTGATGTT	TTGTTTTTCTCAAATTGTAAGT	TTATTTCATTGCGTACTTTA	MAGGATCGCTATAATA
		,**********		+ 800
ACTANTTTTTTGTAGTGGAAAGC	CTAGCTTCCCACTACA	MACAMAAGAGTTTAACATTC	MATAMAGTANCGCATGANAT	TTTTCCTAGCGATATTAT
ikkhhls	d r r v m'f	cfsqivs	ifhevi.	k g s l
. I knitfr		Fvflkl.v		
d.ktspfs		lffsnck		
,				, <del>-</del>
iilfc.re	s r l t i r	ngke.itl ktkrlny:	kn.qts.	rpasyy lfsr.i
h n f f m v k t	rispin odfnet	knkefql	nienry k	fliaii
3 V G 3 K	, , , , ,			
		<b></b> .		BXM
E		ĆH FP		ASA
c		RA		NPE
D		22		223
•				//
ACCANTAAGGACAAATGAATAA	AGATTGTATCCTTCGGG	GCAGGGTGGAAATCCCGACCGG	CGGTAGTAAAGCACATTTGC1	TTAGAGCCCGTGACCCGT
				980
TEGTTATTCCTGTTTACTTATT	TCTAACATAGGAAGCCC	CGTECEACETTTAGGGETGGCE	GCCATCATTTCGTGTAAACG	WAI CI COOPTACI GAGEN
pirtne.	rlypsg	qggnpdr	rstfa	lepvtr
a, gamnk			gskahll	. s p . p v
tnkdk.ik				rard pc
		,		k s a t v r
		pltsigv		. lghgt
		aphfdrg	F	
'' '		_		

FIGURE 3C

					Pag
TD H H A		F	F F	MD	
HRN SL		0	N O	SR	
		ĸ	UK	EA	
1,3 1 1 1		1	H 1	11	
TGCATAAGCACGCGGGTGGATTCAGTTTAAGCT					
CACGTATTCGTGCGCEACCTAAGTCAAATTCGA					1000
hkhavdsv.,acistruiqfkla-arggfsls.	kptvkvuag ( srq.ksgwe	g. kdd	aamq eplek	celkmh nv.kc mfkna	
tcicatset, a					
hmivrhi.nla	fgvtftqip	<b>s</b> p h	haaic	ft.fh	
ay arppniki	qireh fd phs	f s s	s g s h i	inlfa	
A	н .		н н		
v	N N		и и		
A	F REO-INDEPENDENT F		i i		
3	3 TRANSCRIPTION 3		1 1		
/	TERMINATOR			RBS	
TAGTGTTATTTCCTATTGCGTAAAATACCTAA	AGECCEGAATTTTTATAAATTEGGGGET	TTTTTGACG	GTAAATAACAAAAG	AGGGGAGGGAAACAA	
ATCACAATAAAGGATAACGCATTTTATGGATT			,+,		1100
The state of the s	, coulded i About A i i i Abbelli i	WWW.ALTEL	EXTITATION 1	iccccicccit i fit	
s v i s y c v k y t k	prifykfea	fle	v n n k r	a	
v l f pia . n t .		f . r	. itke	<b>aret</b> n	
. cyflirkipk	a on fl. ired	۔ د من		raaka	
		T G G		*****	
,	fgrik. Inna	k k v	t f l l l	D & D f 1	
ltie, qtfyr Tingiay fv.	fgrik. Inpa Igsnkifep:	k k v	t f l l l y i y f	p	
ltie, qtfyr tnngiayfv.	fgrik. Inpa Igsnkifep:	k k v	t f l l l y i y f	p	
ltie, qtfyr Itnngiay fv.	fgrik. Inpa Igsnkifep:	k k v	t f l l l y i y f	p	
Ltie.qtfyr ntnngiayfv. yh.krnrligl	fgrik. Inpa Igsnkifep:	k k v	t f l l l y i y f	p	
Ltie.qtfyratnngiayfv.yh.krnrligl  E M A HHD XBMD  A B L AAD HGBP	fgrik.lnpa lgsnkifeps agfkkyirp	k k v	tflll yivf blycf	p	
Ltie.qtfyratnngiayfv.yh.krnrligl EMAHDXBMD	fgrik.lnpa lgsnkifepa agfkkyirp HEH	k k v	tflll yivf blycf HHH	p	
Ltie.qtfyratnngiayfv.yh.krnrligl E M A HHD XRMD A 3 L AAD HGEP R O U EEE CLON	fgrik.lnpa lgsnkifeps agfkkyirp HEH NCN	k k v	tflll yivf blycf HHH IHA	p	
Ltie.qtfyratnngiayfv.yh.krnrligl EE MAHDXBMD BA BLAADHGBP R O U EEE OLON 21 2 1 131 2211	fgrik.lnpa lgsnkifeps agfkkyirp HEH NCN FPF 311	k k v ; k q r k k s ;	tflll yivf blycf HHH IHA NAE P12	pspfl splsv lppfc	
Ltie.qtfyr ntnngiayfy. yh.krnrligl  EE M A HHD XBMD  AA B L AAD HGBP  PR O U EEE CLOM  21 131 2211  CORF 5 / /// GGAAGAGTATTATATGAAGCTGGCCTTAGAT	fgrik. Inpa Igsnkifep: agfkkyirp  HEH NCN FPF 311	kkv kqr kks	tflll yivf blycf HHH IHA NAE P12	pspflsv lppfc	
Ltie.qtfyr ntnngiayfy. yh.krnrligl  KE M A HHD XBMD  BA B L AAD HGBP  PR O U EEE CLOM  21 131 2211  /_ORF 5 / ///  GGAAGAGTATTATATGAAGCTGGCCTTAGAT	fgrik.lnpa lgsnkifeps agfkkyirp  HEH NCN FPF 311	k k v s k q r k k s j	tflll yivf blycf HHH IHA NAE P12	p s p f l s p l s v l p p f c	1200
Ltie.qtfyrmtnngiayfy.yh.krnrligl KE MAHD XBMD SA 3 L AAD HGEP PR 0 U EEE CLOM 21 21 131 2211 /_ORF 5 / /// GGAAGAGTATTATATGAAGCTGGCCTTAGAT	fgrik.lnpa lgsnkifeps agfkkyirp  HEH NCN FPF 311	k k v s k q r k k s j	tflll yivf blycf HHH IHA NAE P12	p s p f l s p l s v l p p f c	1200
L t i e . q t f y r m t n n g i a y f v . y h . k r n r l i g l  KE M A HHD XRMD SA B L AAD HGEP PR O U EEE CLON 2 1 131 Z211 // ORF 5 /// FGGAAGAGTATTATATGAAGCTGGCCTTAGAT	f g r i k . L n p a L g s n k i f e p s a g f k k y i r p  HEH NCN FPF 311  CTTGCGAAGCAGGGCGAAGGACAGACCGA	k k v s k q r k k s s	t f l l l y i v f b l y c f  H HH I HA N AE P 12  GCTCGTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	p s p f l s p l s v l p p f c	1200
L tie. q t f y r m t n n g i a y f v . y h . k r n r l i g l  KE M A HHD XBMD SA B L AAD HGBP PR O U EEE OLON 21 2 1 131 2211  /_ORF 5 / /// GGAAGAGTATTATATGAAGCTGGCCTTAGAT  ACCTTCTCATAATATATCTTCGACCGGAATCTAG  K S i i S W p . i	f g r i k . l n p a l g s n k i f e p s a g f k k y i r p  HEH NCN FPF 311  CTTGCGAAGCAGGGCGAAGGACAGACCGA	k k v s k q r k k s s	t f l l l y i v f b l y c f  H HH I HA N AE P 12  CCTCGTCGGCGCTG  CGAGCAGCCGCGAC	p s p f l s p l s v l p p f c  TTGTCGTAMAGGACG  MACAGCATTTCCTGC	1200
Ltie.qtfyrmtnngiayfv.yh.krnrligt  KE M A HHD XEMD  SA B L AAD HGEP  PR O U EEE OLON  21 2 1 131 2211  /_ORF5 / ///  IGGAAGAGTATTATATGAAGCTGGCCTTAGAT  ACCTTCTCATAATATACTTCGACCGGAATCTAG  K S i . S W p . i  g r v l y e a g i r s	f g r i k . l n p a l g s n k i f e p s a g f k k y i r p  HEH NCN FPF 311  CTTGCGAAGCAGGGCGAAGGACAGACCGA  CAACGCTTCGTCCCGCTTCCTGTCTGGCT	k k v s k q r k k s j	t f l l l y i v f b l y c f  H HH I HA N AE P 12  ICTCGTCGGCGCGAC	p s p f l s p l s v l p p f c  TTGTCGTAAAGGACG  AACAGCATTTCCTGC	1200
Ltie.qtfyr ntnngiayfv. yh.krnrligt  KE M A HHD XEMD  SA 3 L AAD HGEP PR O U EEE OLON 21 2 1 131 2211  /_ORF 5 / /// IGGAAGAGTATTATATGAAGCTGGCCTTAGAT  ICCTTCTCATAATATACTTCGACCGGAATCTAG  K s i i . s w p . i g r v l y e a g l r s e e y y m k l a l d !	fgrik.lnpa lgsnkifeps agfkkyirp  HEH NCN FPF 311  CTTGCGAAGCAGGGCGAAGGACAGACCGAT  CAACGCTTCGTCCCGCTTCCTGTCTGGCT  Lrsrakdrpn ceagrrtdr lakqgegqte	k k v s k q r k k s j	tflll yivf blycf HHH IHA NAE P12 GCTCGTCGGCGCGGGGCGGGGGGGGGGGGGGGGGGGGG	p s p f i s p l s v l p p f c	1200
Ltie.qtfyrmtnngiayfy. yh.krnrligl  KE M A HHD XBMD  SA 3 L AAD HGEP  PR O U EEE CLON  21 2 1 131 2211  /_ORF5 ////  GGAAGAGTATTATATGAAGCTGGCCTTAGAT  ACCTTCTCATAATATATCTTCGACCGGAATCTAG  k s i . s w p . i g r v l y e a g i r s e e y y m k l a l d i	fgrik.lnpa lgsnkifep: agfkkyirp  HEN NCN FPF 311  TTTGCGAAGCAGGGCGAAGGACAGACCGA  CAACGCTTCGTCCCGCTTCCTGTCTGGCT  Lrsrakdrpn ceagrrtdr	k k v s k q r k k s s	tflll yivf blycf HHH IHA NAE P12 GCTCGTCGGCGCTG GCAGCAGCCGCGAC	P & P f l s P l s v l P P f c  TTGTCGTAAAGGACG  AACAGCATTTCCTGC l s - r t c r k g r v v k d g	1200
Ltie.qtfyrntnngiayfv.yh.krnrligl  KE MAHD XBMD  SA BLAAD HGBP  PR DU EEE CLOM  21 131 2211  /_ORF5 ///  IGGAAGAGTATTATATGAAGCTGGCCTTAGAT  ACCTTCTCATAATATATTCTCGACCGGAATCTAG  k s i i . s w p . i g r v l y e a g l r s e e y y m k l a l d !  h f l i i h l q g . i	fgrik.lnpa lgsnkifeps agfkkyirp  HEN NCN FPF 311  CTTGCGAAGCAGGGCGAAGGACAGACCGA  GAACGCTTCGTCCCGCTTCCTGTCTGGCT lrsrakdrpn ceagrrtdr lakqgegqte	kkv kqrkks qr	tflll yivf blycf HHH IHA NAE P12 CCTCGTCGGCGCTG	P S P f l S P l S V l P P f c  TTGTCGTAAAGGACG  AACAGCATTTCCTGC l S	1200
Ltie.qtfyrntnngiayfy. yh.krnrligl XE M A HHD XBMD AA B L AAD HGEP PR O U EEE CLON 21 2 1 131 2211 /_ORF 5 / /// GGAAGAGTATTATATGAAGCTGGCCTTAGAT  CCCTTCTCATAATATACTTCGACCGGAATCTAG I k s i	fgrik.lnpa lgsnkifeps agfkkyirp  HEH NCN FPF 311  CTTGCGAAGCAGGGCGAAGGACAGACCGAT  GAACGCTTCGTCCCGCTTCCTGTCTGGCT  lrsrakdrpn ceagrrtdr lakqgegqte krllafslg	kky k q r k k s ;  ATCCAATCCC  TAGGTTAGGC	tflll yivf blycf HHH IHA NAE P12 CCTCGTCGGCGCTG	P S P f l S P l S V l P P f c  TTGTCGTAAAGGACG  AACAGCATTTCCTGC l S	1200

FIGURE 3D

								Pag
	ON HHEHIAANS		A HNF	N	N <b>S</b> B	HHM NN	BN	
	AAHLABHI SR		L PLO	L	DS S	NGS SL	AL	
	NRAAEEAN EA		U HAK	A	ET T	LIP PA	N A	
	1124211P 11		1 131	3	1X X	1A2 H3	1 4	
	// ///		//			/ /		
CAAATTGTCG	GAATGGGCGCCCATTTAA	AATATGGTGA	GCTCATGCAGAAG	TTEATGECAT	CCATATGGCT	GAGCACATGCAG	AGGGTGCCGACAT	
····		,						130
TTTAACAGC	CTTACCCGCGGGTAAATT	TTATACCACT	CGAGTACGTCTTC	AAGTACGGTA	GGTATACCGA	ECTEGTGTACGTE	TCCCACGGCTGTA	
kls	ewapi.	nt m v k	lmqk	fmps	iul	e h m q	rvptf	
пег	ngrpfk	i w . :		s c h	p y g w	ster	gcrh	
a i v g	ngah i k	y g •	a h a e v	hai	h m a	g	gadi	
	,							
lndi	shagm.	fit	fsmcf	ппд	dmhs	\$ C M C	ltgv	
y f q r	fprgnl	, i h h	iehli	. e a w	9 Y P	qivhi	phre	
CIT	pipawk	турѕ	• . • •			p	• p • s 41	
IBM 7		В	N				X	
A AZ		В	L				HE	
EAE C		V	A				00	
213 1	H	1	3				21	
/	CGAACCGTGCAGCCATT						-	//
	GCTTGGCACGTCGGTAAT					AGTTTTCTCACAA		146
t	snraaî 1 rtvqpl	rgcctttttgt tekh rknt	F h V, q :	TAACTAATAC	TTGAGACCAT t l v t l w y	AGTTTTCTCACAA s k e c s q k s v	GCACCGCTACTCT	146
t	enraait rtvqpl epcshy	recentitier tekh rknt gkt	F h V, Q i a m c r p p c a e	TACTATAC i d y c l i i	t l v t l v t l w y n s g i	AGTTTTCTCACAA skecs qksv krvf	GCACCGCTACTCT  H	140
t   h s ryt v t	snraai 1 rtv qpl epcshy	recentification in the second	rhy,qa amcr ppcae	TAACTAATAG idyg lij	t l v t l v t l w y n s g i	AGTTTTCTCACAA S k e c s q k s v k r v f	GCACCGCTACTCT  H	140
t l h s ry t v t l	snraaî 1 rtvqpl epcshy	recentities  tekh  rknt  gkt   vsfc	FREE C	TAACTATAC idyc lii fqnc	tlv tlv ilwy nsgi	AGTTTTCTCACAA skecs qksv krvf 	GCACCGCTACTCT  Wr.e rgder vamr  ehrhs	140
t l h s r y t v t l	snraaii rtvqpl epcshy efraam rvtcgn	racctttttat tekh rknt gkt: vsfc rff	COCCOTACACGTC  Thy, qi amcr ppcae  rwtc vamhi	TAACTATAC idyo Lii fqno is.	tiv tiv qiwy nsgi 	AGTTTTCTCACAA skecs qksv krvf dfsh	AGCACCGCTACTCT  SHIP.e  Fgder  Vamr  ehrhs  Fpss	146
t l h s r y t v t l	snraaî 1 rtvqpl epcshy	racctttttat tekh rknt gkt: vsfc rff	COCCOTACACGTC  Thy, qi amcr ppcae  rwtc vamhi	TAACTATAC idyo Lii fqno is.	tiv tiv qiwy nsgi 	AGTTTTCTCACAA skecs qksv krvf dfsh	AGCACCGCTACTCT  SHIP.e  Fgder  Vamr  ehrhs  Fpss	140
t l h s r y t v t l	eraait rtvqpl epcshy efraam rvtcgn	recentites  tekh rknt gkt:  vsfc rff	Fhv, q a a m c r p p c a e r r r v t c v a m h l g g h a s	TTAACTIATAC  I dy c Lii  f q n c i s	tiv qiwy nsgi dvrt .sqy	AGTTTTCTCACAA skecs qksv krvf dfsh cflt illtr	GCACCGCTACTCT  GM	141
t l h s r y t v t l v n c r . v t v s	ertveplepcshy ertvegn rvtcgn sghlw.	recentities  tekh rknt gkt:  vsfc rff pfv	Fhv, q a a m c r p p c a e r r w t c v a m h l g g h a :	TTAACTIATAC 1 dy c Lii f q n c i s c	tlv qlwynsgi 	AGTTTTCTCACAA  & k e c a q k s v k r v f d f s h i l l t r	GCACCGCTACTCT  SM	141
t l h s r y t v t l v n e r . v s BH	ertveplepcshyeraam rtvepleraam rvtegn sghlw.	recentities  tekh rknt gkt:  vsfc rff pfv	F h V, q a a m c r p p c a e r r w t c v a m h i g g h a :	TTAACTIATAC	tlv qlwynsgi nsgi 	AGTTTTCTCACAA  S k e c s q k s v k r v f  d f s h i l l t r  AS S E EAS	AGCACCGCTACTCT  EMF.e  Fgder  Vamr  ehrhs  rpss  tail	141
t lhs ryt vtl vnc r.v	eraaii rtvqpl epcshy efraam rvtcgn sghlw MKE BM NSA IB LPR NO	recentification in the kerne in the policy of the policy o	F h V, q a a m c r p p c a e r r u t c v a m h l g g h a :	TTAACTIATAC  I d y c  L i i  f q n c  i d ,  f q n c  i d ,  f q n c  i d ,	t l v q l w y n s g i d v r t . s q y l e p	AGTTTTCTCACAA  8 k e c s q k s v k r v f d f s h f l t i l l t r  EAS S E EAS	AGCACCGCTACTCT  SHIP e  F g d e r  V a m r  e h r h s  r p s s  h t a i l	140
t lhs ryt vtl vnc r.v	e fraam rvtegn s ghlw  MKE BM NSA IB L PR NO 1 21 11	recentification to the k h r k n t g k t in the control of the con	F h V, q a a m c r p p c a e r r w t c v a m h i g g h a :	TTAACTIATAC	t l v q l w y n s g i d v r t . s q y l e p	AGTTTTCTCACAA  8 k e c s q k s v k r v f d f s h f l t i l l t r  EAS S E EAS PC F C CPE LYR A R PYF 11 N 2 11	AGCACCGCTACTCT  SHIP e  F g d e r  V a m r  e h r h s  r p s s  h t a i l  SH  SH  SH	140
t l h s r y t v t l	rtvqplepcshy efraam rvtcgn sghlw  MKE BM NSA IB L PR NO 1 21 11	TRECETTITGE  TREE	Fhv, qaam cr ppcae rwtc vamht ggha:	TTAACTIATAC  I dy c Lii  T q n c i s s n i i  F O K 1	t l v q l w y n s g i d v r t . s q y l e p	AGTTTTCTCACAA  & k e c s q k s v k r v f ,	AGCACCGCTACTCT  SM	
t l h s r y t v t l v n c r . v t v s	e fraam rvtcgn s ghlw  MKE BM N SA IB L PR NO 1 21 11 / RBS	recentities  to e k h  r k n t  g k t  v s f c  r f f  p f v  D H H  P B L  N O A  1 2 3	F h V, Q a a m c r p p c a e r r w t c v a m h l g g h a :	TTAACTIATAC  I dy c Lii  f q n c i a . s n i i  F C CATTGAGGTA	t i v q i w y n s g i s q y i s q y i s p	AGTTTTCTCACAA  S k e c s q k s v k r v f ,	AGCACCGCTACTCT  S W	
t l h s r y t v t l v r c r v s en iA nM iII /.tcctaatcc	e fraam rvtegn e fraam rvtegn s ghlw MKE BM NSA IB L PR NO 1 21 11 / RBS	recentification to the kind of	F h V, Q is a m C r p p c a e e r v a m h is g g h a :	TTAACTIATAC  I d y c L i i  F Q n c i s . s n i i  CATTGAGGTA	t l v q l w y n s g i	AGTTTTCTCACAA  R k e c s q k s v k r v f  d f s h c f l t i l t r  EAS S E EAS PC F C CPG YR A R PYF 11 N 2 11' //CCTGGCAGACCAG	GCACCGCTACTCT  S W	151
t l h s r y t v t l v n c - t v s	e fraam rvtcgn s ghlw  MKE BM N SA IB L PR NO 1 21 11 / RBS	recentification to the kind of	F h V, Q is a m C r p p c a e e r v a m h is g g h a :	TTAACTIATAC  I d y c L i i  F Q n c i s . s n i i  CATTGAGGTA	t l v q l w y n s g i	AGTTTTCTCACAA  R k e c s q k s v k r v f  d f s h c f l t i l t r  EAS S E EAS PC F C CPG YR A R PYF 11 N 2 11' //CCTGGCAGACCAG	GCACCGCTACTCT  S W	151
t l h s ryt v t l v t l ir. v	e fraam rvtegn e fraam rvtegn s ghlw MKE BM NSA IB L PR NO 1 21 11 / RBS	recentification to the kind of	F h V, Q is a m C r p p c a e e r v a m h is g g h a :	TTAACTIATAC  I d y c L i i  F Q n c i s . s n i i  CATTGAGGTA	t l v q l w y n s g i	AGTTTTCTCACAA  R k e c s q k s v k r v f  d f s h c f l t i l t r  EAS S E EAS PC F C CPG YR A R PYF 11 N 2 11' //CCTGGCAGACCAG	GCACCGCTACTCT  S W	15(
t l h s r y t v t l v n e r v s i n v s	e fraam rvtegn e fraam rvtegn s ghlw MKE BM NSA IB L PR NO 1 21 11 / RBS	TRECETTITIET  THE RESERVENCE	F h V, Q : a m C r p p C a e r t C v a m h t g g h a :  TGAAAGAAGCTGG	TAACTIATAC  I d y c  L i i  f q n c  i d y  f q n c  i d y  CATTGAGGTA  GTAACTICAT	t l v q l w y n s g i	AGTTTTCTCACAA  R k e c s q k s v k r v f c f s h c f l t i l l t r  AS S E EAS PC F C CPE YR A R PYE 11 N 2 11 // CCTGGCAGACCAC	AGCACCGCTACTCT  S M	151
t l h s r y t v t l v n e r v s i n v s i n i n i i n i n i i n i i i i cctaatcci	T T V Q P L  P C S h Y  F T V C G N  S S G h L W  M KE BN  N SA IB  L PR NO  1 21 11  / RBS  GCTTGTGGCTGGAAGAGG  CGAACACCGACCTTCTCC	TRECETTITIET  T R N T	F h V, Q : a m c r p p c a e r t c v a m h i g g h a : AM LN UL 11 TGAAAGAAGCTGG	TTAACTIATAC  I d y c  L i i  f q n c  i s ,  s n i i  CATTGAGGTA  GTAACTICAT	t i v q i w y n s g i	AGTTTTCTCACAA  S k e c s q k s v k r v f  d f s h i l l t r  AS S E EAS PC F C CPE YR A R PYF 211 N 2 11 // CCTGGCAGACCAG	EGCACCGCTACTCT  IN P. e  F. g. d. e  F. g. d. e  F. g. d. e  F. p. s. s.  F. p. s. s.  EMARCH   151	
t l h s r y t v t l v n c r v s in H in H in I in C in GGATTAGGI	e fraam rvtcgn s ghlw  MKE BM NSA IB L PR NO 1 21 11 / RBS GCTTGTGGCTGGAAGAGGG  GGAACACCGACCTTCTCC	TRECETTITIES  The second secon	F h V, Q : a m c r p p c a e  r w t c v a m h i g g h a :  AM LN UL 11 TGAAAGAAGCTGG	TTAACTIATAC  I d y c  L i i  f q n c  i s ,  s n i i  CATTGAGGTA  GTAACTICAT  L r ,  h g k	t i v q i w y n s g i	AGTTTTCTCACAA  S k e c s q k s v k r v f c f t i t t r  AS S E EAS PC F C CPC YR A R PYS 11 N 2 11 / CCTGGCAGACCAG GGACCGTCTGGTG GGACCGTCTGGTG W q t r p g r p s	CCGCCTCTCCGACT	151
t l h s r y t v t l v n c r v s en l in	T T V Q P L  P C S h Y  e f r a a m  r V t c g n  s g h L  N KE BM  N SA IB  L PR NO 1 21 11  / RBS  GCTTGTGGCTGGAAGAGG  CGAACACCGACCTTCTCC  L w L e e g  a c g w k r  L V a g r g	TRECETTITIES  The kent  Th	F h V, Q a a m c r p p c a e r r v c v a m h i g g h a a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m h i g g h a m c r v a m	ITAACTIATAC  I d y c  L i i  f q n c  i d y  f q n c  i d y  CATTGAGGTA  CTAACTICAT  CTAACTICAT  h g k  i e v	t l v q l w y n s g i	AGTTTTCTCACAA  S k e c s q k s v k r v f  d f s h i l l t r  AS S E EAS PC F C CPC YR A R PYS 11 N 2 11' // CCTGGCAGACCAG GGACCGTCTGGTG GGACCGTCTGGTG G w q t r p g r p s l a d q	CCCCCTCTCCGACT  C g g e a e a e r l n	151
t l h s r y t v t l v n c r . v t v s HA IA	e fraam rvtcgn s ghlw  MKE BM NSA IB L PR NO 1 21 11 / RBS GCTTGTGGCTGGAAGAGGG  GGAACACCGACCTTCTCC	TGCCTTTTTGTI  T & h T R N T T S T T T T T T T T T T T T T T T T	COCCETACACETC  The Very Camber	TTAACTIATAC  I d y c L i i  f q n c i s . s n i i  CATTGAGGTA  CAT	t i v q i w y n s g i d v r t . s q y i e p  ACCCAAGGCAT TCCCTTCCGT/	AGTTTTCTCACAA  S k e c s q k s v k r v f c f s h c f l t i l t r AS S E EAS L T R R PYS L T R R R PYS L T R R R R R R R R R R R R R R R R R R	AGCACCGCTACTCT  SHIP E F g d e P V a m P  e h r h s P p s s h t a i l  SH RL H GGCGGAGAGGCTGA  CCGCCTCTCCGACT  P g e a e a e p l n	15
T L h s F Y T V T L V T C F V S  BM IA NM 11 / TCCTAATCC  AGGATTAGGI L i F S S S P D D	T T V Q P L e p c s h y e f r a a m r V t c g n s g h L N KE BN N SA IB L PR NO 1 21 11 / RBS GCTIGTGGCTGGAAGAGG CGAACACCGACCTTCTCC L w L e e g a c g w k r L V a g r g	TECCTTTTTGTI  TERM TERM TERM TERM TERM TERM TERM TERM	F h V, q a a m c r p p c a e r v a m h l g g h a :  AM LN UL 11  TGAAAGAAGCTGG  ACTITCTTCGACC	TAACTATATAC  I d y c  L i i  f q n c  i s  s n i i  CATTGAGGTA  CATTGAGGTA  CATTGAGGTA  CATTGAGGTA  CATTGAGGTA  A n y k  i e v	t i v  i i w  i i w  i i w  i i w  i i s  g  i o  contact  AGGGAAGGGAT  TCCCTTCCGTJ  g  k  g  g  r  e  g  r  e  g  r  e  g  r  f  f  f  f  f  f  f  f  f  f  f  f	AGTTTTCTCACAA  S k e c s q k s v k r v f , d f s h i l l t r  AS S E EAS PC F C CPC YR A R PYI 11 N 2 11' // CCTGGCAGACCAG GGACCGTCTGGTC GGACCGTCTGGTC G w q t r p g r p 1 l a d q d q c v	CCCCCTCTCCACT  C g e a e  C g e a e  C c g e a e  C c c c c c c c c c c c c c c c c c c	15

FIGURE 3E

X M SHH SS RMM B FF B AH M N TAA PP SAA B N N G LG A N L UEE LLAEE V UU L UE E 1 113 11123 1 S N 1 1 22 3  WAMATTICTGCACTITATGAGGACAGGCCTTCCCGTAGGTCAGGCCGCTGCCACGCTTGACGGCCGTCACGGTCACGGTCACGGTCACGGTCACGTCTTATGAGGACAGGCCTTCCCGACGCTTCCACGCTTGACGCCGTCACGCTTCACCGCTCACGTCCACGACGACGCACGC							•				
HENDREN AFS ED F PIBRIBA LIF SD 0 RIGHT SD 0											Page
HENDREN AFS ED F PIEDREN LAF SD 0  HENDREN LAF SD 0  HENDREN LAF SD 0  HONDRUS LAG ST ST S S S S S S S S S S S S S S S S	٠		CHR C	C PMM	R	F	F	8	A H	н	
MAMATHERCAGTTAAGGACAGGCCTTCCGTACGCCAGTCCCAGTCCCAGTCCCAGTCCCAGTCCACAGTCCACAGTCCACAGTCCACAGTCCAGTCCCAGTCCCAGTCCCAGTCCCAGTCCCAGTCCCAGTCCCAGTCCCAGTCCCACGTCCACAGTCCACACTCCCACGTCCACACTCCCACCTTCCACCACTCCACACTCCCACACTCCACACTCCACACTCCACACTCCACACTCCACACTCCACACTCCACACTCCACACTCCCACACTCCCACACTCCACACTCCACACTCCCACACTCCACACTCCACACTCCCACACTCCACACTCCCACACTCCACACCCCACACTCCACACCCCACACTCCACACCCCACACTCCACACCCCACACTCCCACACTCCCCACACTCCCACACTCCCCACACTCCCACACTCCCCACACTCCCACACCCCACACCCCCACACTCCCCACACTCCCACACTCCCCACACACCCCCACACTCCCCACACTCCCCACACTCCCCACACTCCCCACACTCCCCACACTCCCCACACACCCCCACACACCCCCACACACCCCCACACAC									L G	A	
AMANTHETGEACHTTATCAGGACAGGCTTCCGTACGTCACGCTAMAGCGGCTGCAGGCTGCAGGCTGACAGGTGACAGAGGACTTACAGTGACTTCCGCAGTTTCAGTGACTAGCGACTCTCCGACAGTGACAGTGACTTTCCGGACAGTGACAGGACTCTCGCACATTTCAGTGACTACCGCATTCACAGTAGGACTCTCCGCATTTCACTTACCGCAGAGGACTCACAGGACTCTCCGCATTTCACTTACCGCAGACGCTCACAGTGCACAGTGCACATTGACCTGCACATTGACACTACCGCACAGTAGACTACAGGACTCCTGCACATTCACTTTCCGCACAGTGCACAGTGCACAGTGCACAGTAGACTACAGGACTCCACAGTGCACAGTAGACTACAGGACTCCACAGTGCACAGTGCACAGAGGACTACAGGACTCCACAGTGCACAGTAGACTACACAGAGACTCACAGAGAGACTACAGAGACTACAGAGACTACAGAGACTACAG								Ĺ	υE	E	
WAMATTICTGCACTTATGAGGACAGGCCTTCCGTACGCCAATTICGCCCGCCGCAATTICGCCGCAAGGCAAG									1 2	3	
THITTTALAGGACAGGCTTCCGTAGTCAGCTAGCTGCAGCTTGCAGCTTCAGGCTTGCAGCTTCAGGCTGCACAGCTTCTACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCACAGCTTCCACACTTCCCGTTCTCCCCACCGTCCCCACTTCCCCCCCC	1 1			1 123	•	•	•	•	-		
THITTAMAGACGTGAMATACTCCTGTCCGAAGGCATGCGATTTTCGCCCGACGTCGGAACTGCCGTTCTATCGATGGTGGCCCACTGTC  k n f c t l . g q a'f r t s r . k r l p a l t a r . l p a r v t a k is a l y e d r p s v r h a k s g c q p . r q d a y q h g . q k f l h f m r t g l p y v t l k a a a s l d g k i a t s t g d s  f f k q v k h p c a k r v d r . f r s g a k v a l y s g a r t v f i e a s . s s l g e t r . a l l p q w g q r c s l . w c p h c s f n r c k i l v p r g y t v s f a a a l r s p l i a v l v p s l  HEMOMEN AFS ED F PISPABA LWF SD 0  HNOMLVE MAA PE K  1111112 HNN 11 1  TACCTAGTGCAGTCTCCAGAGGTCCTAGGAATACAGGAAAACACACCAAGGATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGAAGGTTCCTCTACGAGTCGTTAGTCCTTTTGTGTGGTTTTCGTAAAATCAGGCTCAACCCTCAAGTGCAAGCCGACAAT  TA s g s r q r l q d r m l s n t s k h t k a f . s e l a q . k p t i m d h v r g c k t g c s a i q e n t p k h f s r s w h s e s r q s w i t s e a a r q d a q q y r k t h q s i l v g v g t v k a d n  f p d r . l s c s l i s l l v p f c v l a n . d s n a c h f g v i i s . t l p q l v p h e a i c s f v g f c k l r l q c l s l r c l h i v d s a a l c s a . c y l f v c w l m k t p t p v t f a s l  AM P B H M CH R TMH MD H N L S S S H A FP S ANN SD G 5 5 1 1 1 1 3 3 22 1 1 111 21 1 1 1 1 1 1 1		ettt.te.ce.c	// ************	CCTACC	TCACG	CTAAAAGCGG	CTGCC	IGCCTTGACGGCA	AGATAGETACCA	GCACGGGTGACAG	
ETHTHAMAGAGGTGAMATACTCCTGTCCGGAAGGCATGCAGTTCCCCCTACTCCCTTTTCCCTCACTGCTCCTACTCCTTTCCCTCACTGCTCCTACTCCCTTTTCCCCCCACTGCCCTTTTTCCCCCCCC	MODEL I I CI GOA	LIIIAIGAGGAC									1600
kisalyedrpsvrhaksgcqp.rqdsyqhg.qkflhfsrtglpyvtlkaaasldgkiatstgds ffkqvkhpcskrvdr.frsgakvalysgartv fieas.sslgetr.allpqwgqrcsl.wcphc sfnrckilvprgytvsfaaslrspliavlvpsl  HBMDMBM AFS ED F PISPMBA LWF SD O HMOMIVE UMA PE K 1111112 NHH 11 1 /// ATGCTACGTCAGGTGCAGACAGGATGCTCAGGAATACAGGAAAACACCAAAGCATTTAGTCGGAGTTGGAAAGCCGACAAT TACCTAGTGCAGGTCTCAGGATGCTCAGGATAGTCGTTTTTTTT	TTTTTAAAGACGT	GAAATACTCCTG	TEEGGAAG	GCATGO	AGTGC	GATTTTCGC	CACCO.	CEGAACTECEST	TETATEGATGGT	CGTGCCCACTGTC	
fieas.sslgetr.allpqwgqrcsl.wcphc sfnrckilvprgytvsfaaalrspliavlvps(  HBMDMBN AFS ED F PIBPBBA LNF SD O  NHOMMLVE MAA PE K  1111112 NHN 11 1 1 1  /// //AGGATCACGTCAAGACAGGATGCTCAGCAATACAGGAAAACACACCAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGCAGGATGCTCAGCAATACAGGAAAACACACCAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGACGTCCTCTCCCCAGGATGATGTCGCTTTTTGGTGTGTTCGTAAAATCAGCCTCAACCGTGCACATTCGGCTGTTA  Ogsrqrlqdrmlsntgkhtkaf.selaq.kpti adhvrgcktgcaafgcaagatgctaagatgtatgtcctttttgtgtgtgttcGTAAAATCAGCCTCAACCGTGCACATTCGGCTGTTA  Ogsrqrlqdrmlsntgkhtkaf.selaq.kpti adhvrgcktgcaagatgtatgtatgtatgtcctttttgtgttgtg	knfet kisa ekflh	l.gq lyed fmrt	a'f rps glp	r t v r	s r ha	.kr ksg lka	l p c q	alta p, rq sldgk	r, lp dsyq ists	arvta hg.q tgds	
HENDMEN AFS ED F PIEPWEA LMF SD 0 HINDRIVE MUA PE K 11111112 NHM 11 1  /// ATGGGTGCAGGGTGGCAGGAGGGTGCTCAGGATAGAGGAAAACAGCCAAAAGCCGATATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGACGTCTCTCACGAGTCGTTAGTCCTTTTGTGTGTG	ffka	v k h p	c a k	r v	d i	fr	s g	a k v a	l y s g	a r t v	
HENDMAN AFS ED F PIBPNBA LNF SD 0 NNOMLVE MJA PE K 1111112 NHN 11 1  /// //AGGATAGGGCTGCAAGACAGGATGCTCAGGAATACAGGAAAACACCAAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT AGGATACAGGTCCCCAAGGACAGGATGCTCAGGATATACAGGAAAACACCAAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT AGGATACAGGTCCCCCAAGGCTCCTACGAGTCGTTAGTCCTTTTGTGTGGTTTCGTAAAATCAGGCCTCAACCGGTCACCAGTGAAAGCCGACAAT  A g s r q r l q d r m l s n t g k h t k a f . s e l a q . k p t i m d h v r g c k t g c s a i q e n t p k h f s r s w h s e s r q s w i t s e a a r q d a q q y r k t h q s i l v g v g t v k a d n  f p d r . l s c s l i s l l v p f c v l a n . d s n a c h f g v i i s . t l p q l v p h e a i c s f v g f c k l r l q c l s l r c l h i v d s a a l c s a . c y l f v c w l m k t p t p v t f a s l  A M P B M CH R TMM MD H M LS S S M A FP S ANN SD G B 1 L T P F E RA A GLF TE A 0 1 1 1 1 3 3 22 1 1111 21  CAGGTTAACCTGGAGACTGCCGAATGTAACAAAAACAGCCCGGTCCGGGTCATACTTGTACCGTACTCTCGATTCCTGAGGACCCTAAAGTGATTTGCC  a . p a d c r m . q n s r f g s y l i p y s r f l r t l k . f a e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  r a . g a s q r i y c f l r n p d y k i g y e r n r l v s f h n a s s l r c v s s h l v a p e p . v y y r v r s e q p r . l s k r			1 6 6		• .	a   1   1	o a		. s l . v	cphc	
HEHDMEN AFS ED F PIEPPHRA LNF SD 0 NHONLUE ULLA PE K 1111112 NHN 11 1  // / ATAGGATCAGGTCAGAGAGAGGATGCTCAGGAATACAGGAAAACACACCAAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGACGTTCTGCTACGAGTCGTTATGTCCTTTTGTGTGGTTTCGTAAAATCAGCCTCAACCGTGCACAGTGAAAGCCGACAAT  A g s r q r l q d r m l s n t g k h t k a f . s e l a q . k p t i m d h v r g c k t g c s a i q e n t p k h f s r s w h s e s r q s w i t s e a a r q d a q q y r k t h q s i l v g v g t v k a d n  f p d r . l s e s l i s l l v p f c v l a n . d s n a c h f g v i i s . t l p q l v p h e a i c s f v g f c k l r l q c l s l r c l h i v d s a a l c s a . c y l f v c w l m k t p t p v t f a s i  A M P B H M CH R THM MD H H L S S S H A F P S ANN SD G S U E T P F E RA A Q Q F T E A O 11 1 1 3 3 22 1 1111 21 1 1  CAGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGGATACCGTATGGCATTCCTGAGACGCCTAAAGTGATTTGGG  a . p a d c r m . q n s r f g s y l i p y s r f l r t l k . f a e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  r a . g a s q r i y e f l r n p d y k i g y e r n r l v s f h n a s s l r c v a s h l l v a p e p . v y y r v r s e q p r . l s k r					+ ·			l r s D	l i a v	lvpsl	
PIBPNBA LNF SD 0 HNONLUE MAA PE K 1111112 NHN 11 1  // / ATGGATCACGTCAGAGGCTGCAAGACAGGATGCTCAGGAAATACAGGAAAACACCAAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT TACCTAGTGCAGTCTCCGACGTTCTGTCCTAGAGTCGTTAGTCCTTTTGTGTGTTTCGTTTAGAACAGCCTCAACGTGAAAGCCGACAAT  A g s r q r i q d r m i s n r g k h r k a f . s e i a q . k p r i m d h v r g c k r g c s a i q e n r p k h f s r s w h s e s r q s w i r s e a a r q d a q q y r k r h q s i l v g v g r v k a d n  f p d r . l s c s l i s i l v p f c v l a n . d s n a c h f g v i i s . r l p q l v p h e a i c s f v g f c k l r l q c i s l r c l h i v d s a a i c s a . c y l f v c w l m k r p r p v r f a s l  A M P B H M CH R TMM MD H H L S S S H A FP S ANN SD G 5 UE T P F E RA A QLF TE A 0 11 1 1 3 3 222 1 1 111 21 1 1  CAGGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTCGATGCGCTAAAGTGATTTGCG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACTCCTGGGATTTCACTAAACGC  a . p a d c r m . q n s r f g s y l i p y s r f l r r l k . f a e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  c s l r c v a s h l v a p e p . v q y r v r s e q p r . l s k r	5 1 11 1 1		μ.	• /	•	• • -		•			
PIBPNBA LNF SD 0 HNONLUE MIA PE K  1111112 NHN 11 1  // / ATGGATCACGTCAGAGGCTGCAAGACAGGATGCTCAGGAATACAGGAAAACACCAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGACGTTCTGTCCTAGCAGTCGTTAGTCCTTTTGTGGTTTCGTAAAATCAGCCTCAACGGTCACTTTCGGCTGTTA  A g s r q r i q d r m i s n r g k h r k a f . s e i a q . k p r i m d h v r g c k r g c s a i q e n r p k h f s r s w h s e s r q s w i r s e a a r q d a q q y r k r h q s i l v g v g r v k a d n  f p d r . l s c s l i s i l v p f c v l a n . d s n a c h f g v i i s . r l p q l v p h e a i c s f v g f c k l r l q c i s l r c l h i v d s a a i c s a . c y l f v c w l m k r p r p v r f a s i  A M P B H M CH R TMH MD N N L S S S N A FP S ANN SD G B U E T P F E RA A QLF TE A 0 1 1 1 1 3 3 222 1 1 111 21 1 1  CAGGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCCGGTTCGGGTCATACTTGATACCGTACTGTCGTGTGGGACCCCTAAAGTGATTTGCG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCCAGTATGAACTATGGCATGCAGGACTCCTGGGATTTCACTAAACGC  a p a d c r m . q n s r f g s y l i p y s r f l r t l k . f a e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  c s l r c v a s h l v a p e p . v q y r v r s e q p r . l s k r			•								
PIBPNBA LNF SD 0  HMDMLVE MAN PE K  1111112 NHN 11 1  // /  ATGGATCAGGTGCAGAGGATGGTCAGGAATACAGGAAAACACCAAAGCATTTTAGTGGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTGC	HEMDMEM	AFS	E	D	F						
HHOMLUE ULA PE K  1111112 NHN 11 1 1  // ATGGATCACGTCAGAGGCTGCAGAGACGGATGCTCAGCAATACAGGAAAACACCAAGCCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGACGGTTCTGTCCTACGAGTCGTTATGTCCTTTTGTGTGGTTTCGTAAAATCAGCCTCAACCGTGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGTCTCCGACGGTTCTGTCCTACGAGTCGTTATGTCCTTTTGTGTGGTTTCGTAAAATCAGCCTCAACCGTGTCACCTTTCGGCTGTTA  TACCTAGTGCAGTCTCCGACGGTTCTGTCCTACGAGTCGTTATGTCCTTTTGTGTGGTTTCGTAAAATCAGCCTCAACCGTGTCAACTTTCGGCTGTTA  TACCTAGTGCAGTCTCCGACGGTTCTGCGTTATGTCCTTTTGTGTGGTTTCGGAAATCAGCCTCAACCGTGCACGTTCGGGTCAAATCAACGCCGATTCGGGTCAAAATCAACCGCTATGTCGGGTCATACCTTGATACGGTAAGGACGCTAAAGTGATTTGCG  AM PBH M CH R THM MD H M  LS SS N A FP S ANN SD G B  UE T PF E RA A QLF TE A O  11 1 1 3 3 22 1 111 21 1 1  CAGCCTTAACCTGCAGACTGCCCGAATGTAACAAAACAGCCCGTTCGGGTCATAGCTTGATACGGTAACGTCTCGAGACGCCTAAAGTGATTTGCG  CTCCAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCCAGTATGAACTATGGCCATGGAGAGTCATTCCTGAACGCC  a. pad c r m. q n s r f g s y l i p y s r f l r t l k. f a  e l n l q t a e c n k t a g s g h t. y r t l d s. g r. s d l r  s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  c s l r c v a s h l v a p e p. v q y r v r s e q p r. l s k r			\$	5D	0						
1111112 NHH 11 1  // /  ATGGATCACGTCAGGGCTGCAAGACAGGATGCTCAGCAAATACAGGAAAACACACCAAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACAAT  TACCTAGTGCAGGCTCCGACGTTCTGTCCTACGAGTGTTATGTCCTTTTGTGTGGTTTTTTTT					ĸ						
AM PBH M CH R TMM MD H M LS SSN A FP SAMN SD G B LS SSN A FP SAMN SD G B LS T PFE RA A QLF TE A O LL T PFE RA A QLF TE A O LL T PFE RA A QLF TE A O LL T PFE RA A QLF TE A O LL T PFE RA A QLF TE A O LL T PFE RA A QLF TE A O LL T PFE RA A QLF TE A O LL T L L L L L L L L L L L L L L L L											
ATGGATCAGGGCTGCAGGCTGCAGGATGCTCAGCAATACAGGAAAACACCAAGCATTTTAGTCGGAGTTGGCACAGTGAAAGCCGACATTTAGTCGGAGTTGCGAGGTGGAAAGCCGACATTTAGTCCTGTTAGTGCGAGTTGGGAGTTGGCACAGTGAAAGCCGACTTTCGGGTGTTAAATCAGGCCTCAACGGGTGACACTTTCGGGTGTTAAATCAGGCCTCAACGGGTGACACTTTCGGGTGTTAAATCAGGCCTCAACGGGTGACACTTTCGGGTGTTAAATCAGGCCTCAACGGGTGGAAAGCCGGTTAGTGCGTTAAATCAGGCCTCAACGGGGTGACACTTTCGGGTGTTAAATCAGGCCTCAAAGCCGACGCTAAACGGGACGCTAAACGGACGG		PATE N			•						
TACCTAGTGCAGTCTCCGACGTTCTGCCTACGAGTCGTTATGTCCTTTTGTGTGTTTCGTAAAATCAGCCTCAACCGTGTCACTTTCGGCTGTTA  ogsrqriqdrmlsntgkhtkaf.selaq.kpti mdhvrgcktgcsaiqentpkhfsrsmhsesrqs witseaarqdaqqyrkthqsilvgvgtvkadn  fpdr.lscslislivpfcvlan.dsnachfgvi is.tlpqlvpheaicsfvgfcklriqcislrc lhivdsaaicsa.cylfvcwlmktptpvtfasi  AM PBH M CH R TMM MD H M LS SSHA FP SAMM SD G B UE TPFE RA A QLF TE A O 11 1 13 3 22 1 111 21 1  CACGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTCGAGTCCCTGAGGACTCCTGCGATTTCACTAAACGC  CTCGAATTGGACGTCTTGACGGCTTACATTGTTTTGTCGGCCCAAGCCCCAGTATGAACTATGGCATGAGAGCTAAGGACTCCTGCGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrflrtlk.fa elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriycflrnpdykigyernrlvsfhna sslrcvashllvapep.vqyrvrseqppr.lskr	// /		CACCATC	TCAGE	BATAC	AGGAAAACAC	ACCAM	GCATTTTAGTEG	AGTTGGCACAG	TAADAGCCGACAAT	1
TACCTAGTGCAGGTTCCTGCCCTACGAGTCGTTAGTCCTTTTGTGTGGTTTCGTAAAATCAGGCTTAAACTCAGGCTTAAAATCAGCCTCAAAGTCGTTTCGTGTTTTGTGTGGTTTCGTAAAATCAGCCTCAAAGTCAGTTCCGAGTTCCTGAGTCGTTCAAAACAACAACAACAACCCCGGTTCGGGTCATACCTGAACTCATGGCATTCCGGGTTCACCTCAAAGTCATTTCGCG  a. padcrm.qnsrfgsylipysrflrtlk.fa  sltcrlpnvtkqpvrvilds.gr.scllpnvtkqpvrvilds.gr.sdlr  sltcrlpnvtkqpvrvilds.gr.scllpnvtkqpvrvilds.gr.sdlr  cacgattaacctgcaactccccaatgtaacaaaaacaacccagttcgagtcatactttgattaccgtaactatcctgaattcctgaactcctgcaatttcactaaaacaaccccattaaactaacaacaacaa	MIGONICALDIC	I GADGE I DESCRIP	10000101								+ 170
adhvrgektgesaiqentpkhfsrswhsesrqs witseaarqdaqqyrkthqsilvgvgtvkadn  fpdr.lseslisllvpfcvlan.dsnachfgvi is.tlpqlvpheaicsfvgfcklrlqelslrc lhivdsaalcsa.cylfvcwlmktptpvtfasl  AM PBHM CH R TMM MD H H LS SSN A FP SAMN SD G B UE TPFE RA A QLFTE A O 11 1 13 3 22 1 111 21 1 1 1 1 1 1 1 1 1									TELLECTOTO	.cttteccetctt/	
adhvrgcktgcsaiqentpkhfsrswingerquestruitsearqqqyrkthqsilvgvgtvkadn  fpdr.lscslisilvpfcvlan.dsnachfgvi is.tlpqlvpheaicsfvgfcklriqclslrc lhivdsaalcsa.cylfvcwlmktptpvtfasi  AM PBH M CH R TMM MD H M LS SSH A FP SANN SD G 5  UE TPFE RA A QLFTE A 0 11 11333 22 1 1111 21 1 1 1  CAGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTCGAGGACGCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCCAAGCCCAGTATGAACTATGGCATGAGGACTCCTGGGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrflrtlk.fa elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriyeflrnpdykigyernrlvsfqpr.lskr	TTACCTAGTGCAG	PETECGACGTTC	IGTECTACI	EAGTCG	TIAIG	1001111616	16611	CUI AMAI CAUCA	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		•
adhvrgcktgcsaiqentpkhfsrswingerquestruitsearqqqyrkthqsilvgvgtvkadn  fpdr.lscslisilvpfcvlan.dsnachfgvi is.tlpqlvpheaicsfvgfcklriqclslrc lhivdsaalcsa.cylfvcwlmktptpvtfasi  AM PBH M CH R TMM MD H M LS SSH A FP SANN SD G 5  UE TPFE RA A QLFTE A 0 11 11333 22 1 1111 21 1 1 1  CAGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTCGAGGACGCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCCAAGCCCAGTATGAACTATGGCATGAGGACTCCTGGGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrflrtlk.fa elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriyeflrnpdykigyernrlvsfqpr.lskr											
FP dr. lscslisllvpfcvlan. dsnach fgvi is.tlpqlvpheaicsfvgfcklrlqclslrc lhivdsaalcsa.cylfvcmlmktptpvtfasl  AM PBHM CH R TMM MD H M LS SSNA FP SANN SD G 5 UE TPFE RA A QLFTE A 0 11 1133 3 22 1 111 21 1 1  CAGGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTCCGGGTCATACTTGATACCGTACTCTCGAGGACGCCTAAAGTGATTTGCG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCCGCCCAAGCCCAGTATGAACTATGGCATGAGAGTCATGCGGTTTCACTGAAACGC  a. padcrm. qnsrfgsylipysrflrtl. f. fa elnlqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriyeflrnpdykigyernrlvsfhna eslrcvashllvapep.vqyrvrseqpr.lskr	) g	riqd	r ==		n t	g K n					2
fpdr.lseslisilvpfcvlan.dsnachfgvi is.tlpqlvpheaiesfvgfcklrlqclslrc lhivdsaalesa.cylfvewlmktptpvtfasi  AM PBHM CH R TMM MD H M LS SSHA FP SANN SD G B UE TPFE RA A QLF TE A O 11 1133 3 22 1 111 21 1 1  CAGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTCCGGGTCATACCTGCATGCCGTACCTCCTGAGGACGCCTAAAGTGATTTGCG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCCAAGCCCATATGAACTATGGCATGAGAGGCTAAGGACTCCTGGGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrflrtlk.fa elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriyeflrnpdykigyernrlvsfhna	nd hv	rgck		* •	, , 9				v n t v	kadn	_
fpdr.lscslisilvpfcvlan.dsnachtgvlis.rtlpqlvpheaicsfvgfcklrlqclslrc lhivdsaalcsa.cylfvcwlmktptpvtfasi  AM PBH M CH R TMM MD H M LS SSN A FP S ANN SD G B UE TPFE RA A QLF TE A O 11 11333 22 1111 21 1 1  CAGGETTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTGAGGACGCCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTAGGCATGAGGACGCTAAAGGACTCCTGGGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrfirtlk.fa elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd rs.gasqriyeflrnpdykigyernrlvsfhna	wits	eaar	Q 0 8	9 9	Υ		· Q :				•
AM PBHM CH R THM MD H H LS S S N A FP S ANN SD G 5  UE T PFE RA A QLF TE A 0  11 1 1 3 3 22 1 1 111 21 1 1  CARGETTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTGGGGCCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAAGGACTCCTGGGATTTCACTAAACGC  a. paderm. qn s r f g s y lip y s r f l r t l k. f a e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  c . g a s q r i y c f l r n p d y k i g y e r n r l v s f h n a e s l r c v a s h l v a p e p . v q y r v r s e q p r . l s k r	··						~ 1	. n . d		hfov	i
AM PBHM CH R THM MD H H LS SSN A FP SANN SD G 5  UE TPFE RA A QLF TE A O  11 1 1 3 3 22 1 1111 21 1 1  CARGETTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTGGAGTCCTGAGGACGCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACGCTAAAGTGATTTCACTAAACGC  a. paderm. qnsrfgsylipysrflrtlk.fa elnlqtaeenktagsght.yrtlds.gr.sdlr slterlpnvtkqpvrvildtvlsipedakvied  ra.gasqriyeflrnpdykigyernrlvsfhna eslrevashllvapep.vqyrvrseqpr.lskr	tpar		5 L 1	* '	•				1001	slrc	
A M P B H M CH R THM MD H M  LS S S N A FP S ANN SD G 5  UE T P F E RA A QLF TE A 0  11 1 1 3 3 22 1 1111 21 1 1 1  CAGGETTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTAGTCTCTGAGGACGCTAAAGTGATTTGCG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACGCTAAAGTGATTTGCG  a . p a d c r m . q n s r f g s y l i p y s r f l r t l k . f a  e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r  s l t c r l p n v t k q p v r v i l d t v l s i p e d a k v i c d  r a . g a s q r i y c f l r n p d y k i g y e r n r l v s f h n a  s l r c v a s h l v a p e p . v q y r v r s e q p r . l s k r	( <b>5</b>	. pq.	V P n		1		•	, <u> </u>		r f a s l	
LS SSHA FP SAHN SD G 5  UE TPFE RA A QLF TE A 0  11 1133 3 22 1 111 21 1  CAGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTCGAGGACGCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGCATGAGGAAGGA	hivd		c s	• .	C Y	t T V	E #		, , ,	• • • • •	
LS SSHA FP SANN SD G S  UE TPFE RA A QLF TE A O  11 11333 22 1 111 21 1  CAGCTTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTAGTCTCGAGGACGCTAAAGTGATTTGGG  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACGCTAAAGTGATTTCAGG  a. padcrm.qnsrfgsylipysrflrtlk.fa  elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriyeflrnpdykigyernrlvsfhna  sslrcvashllvapep.vqyrvrseqpr.lskr								·			
LS SSHA FP SANN SD G B  UE TPFE RA A QLF TE A O  1 1 1 1 3 3 22 1 1 111 21 1 1  AGCITTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTCGAGTCATGGGACGCTAAAGTGATTTGGG  TCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACGCTAAAGTGATTTGGG  18  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACTCCTGCGATTTCACTAAACGC  a. padcrm. qnsrfgsylipysrflrtk.fa  elnlqtaecnktagsght.yrtlds.gr.sdlr  sltcrlpnvtkqpvrvildtvlsipedakvicd  ra. gasqriyeflrnpdykigyernrlvsfhna  sslrcvashllvapep.vqyrvrseqpr.lskr				CH				R THH	MD	н н	
UE TPFE RA A QLF TE A O 11 1133 3 22 1 1111 21 1 1  AGGETTAACCTGCAGGCTGCGGATGTAACAAAACAGCCGGTTCGGGGTCATACTTGATACCGTACTCTGGAGGACGCTAAAGTGATTTGCG  TCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTGATGAGGATGAGGACTCCTGGAGTCATGAGTGATTTGCG  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAAGCCTATGACTGATGAGGATTCCTGAGGACGCTAAAGTGATTTGCG  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCAAGCCCAAGCTATGACTATGACCTAAACGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCAAGCCCAAGGACTAAAGGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCAAGCCCAAGGACTCCTGGATTTCACTAAACGC  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCAAGCCCAAGGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCAAGCCCAAGCTAAAGGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCCAAGCCCAAGCTAAACGGCTAAAGTGATTTGCG  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCCAAGCCCAAGCTAAACGGCTAAAGGGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTTGTCGGCCAAGCCCAAGCTAAACGGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTTACATTGTTTTTGTCGGCCAAGCCCCAAGCCTAAACGGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGTTACATTGTTTTTGTCGGCCAAGCCCCAAGCCTAAACGTGATTGACTAAGCGACTCCTGGGATTTCACTAAACGC  18  TCGAATTGGACGCTTACATTGTTTTTGTCGGCCAAGCCCCAAGCCTAAACGGACTCCTGGATTTCACTAAACGCC  18  TCGAATTGGACGTTACATTGTTTTTTTTTTTTTTTTTT										G 5	
11 1133 22 1 1111 21 1 1 1 1 1 1 1 1 1 1										_	į
CAGETTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTTGATACCGTACTCTGGATTCCTGAGGACGCTAAAGTGATTTGCG  18  CTCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGGCATGAGGACTCCTGGGATTCACTAAACGC  a. paderm. qnsrfgsylipysrflrtlk.fa elnlqtaecnktagsght.yrtlds.gr.sdlr slterlpnvtkqpvrvildtvlsipedakvicd  c. lerlpnvtkqpvrvildtvlsipedakvicd  ra. gasqriyeflrnpdykigyernrlvsfhna								• • • • • • • • • • • • • • • • • • • •			
AGETTAACCTGCAGACTGCCGAATGTAACAAAACAGCCGGTTCGGGTCATACTGATACCGTACTCTCGAGTCATCCTGAGGACGCTAAAGTGATTTGCG  18  TCGAATTGGACGTCTGACGGCTTACATTGTTTTGTCGGCCAAGCCCAGTATGAACTATGGCATGAGGACGCTAAAGGACTCCTGGGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrflrtlk.fa elnlqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd  ra.gasqriyeflrnpdykigyernrlvsfhna eslrcvashllvapep.vqyrvrseqpr.lskr	1 1	113 3		22						•	
TEGANTEGACGECTTACATTGTTTTTGTCCGCCAAGCCCAGTATGACTATGGCATGAGGAGGCTAAGGACTCCTGCGATTTCACTAAACGC  a.padcrm.qnsrfgsylipysrflrtlk.fa elnlqtaecnktagsght.yrtlds.gr.sdlr slterlpnvtkqpvrvildtvlsipedakvicd ra.gasqriyeflrnpdykigyernrlvsfhna sslrevashllvapep.vqyrvrseqpr.lskr										**********	·n
a.padcrm.qnsrfgsylipysrflrtlk.fa elnlqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd	CAGCTTAACCTGC	AGACTGCCGAAT	GTAACAAA	ACAGC	256770	GGGTEATAC	TIGATA	CCGIACICICIA	I CL I GABGACGE		. 18
a.padcrm.qnsrfgsylipysrflrtlk.fa elniqtaecnktagsght.yrtlds.gr.sdlr sltcrlpnvtkqpvrvildtvlsipedakvicd 	••						• • • • • • • • • • • • • • • • • • •				- 10
e interestate of the state of t	CTCGAATTGGACG	TCTGACGGCTTA	CATTETTT	TGTCG	CCAA	CCCAGTATG	AACTAT	GECATGAGAGETA	AGGACTECTGE	NI I I CAC I AAAC G	16
e interestate of the state of t											
e interestate of the state of t	а.ра	derm	. q n		r f	g s y	( i	p y	trt	, k	
s tertpnytk qpyrvildtyl sipedak vied 	elnic	: t a e c	n k	t a	G 5	ght	. Y	rtla	8 . g r	. 2 0 1	•
ra.gasqriyoflrnpdykigyernrlvsfhna estrovashttvapep.vqyrvrseqpr.tskr		- 1 - 5	vtk	a b	V (	- v i l	<b>a</b> t		P 4 G 8	K V 1 C	•
ragas qriyoflen pdykigyern elvsth na celecvash llvapep, v qyevese qpe, lske							*****	••• , ••••••			•
estrovashit vapep. vqyrvrseqpr. iskr	ra. 9 a	sari	y e	fl	r n	рфу	k i	gyer	nriv	sthia	•
it was a fry feat remark the idssalt id	6	C V 8 S	hll	V 8	<b>p</b> 1	e p . v	' Q )		eqpr	. isk	٢
					o T	rtm	8 5	vtsei	g 5 5 i	altiq	

FIGURE 3F

H HH H H THINK  I HA L HINH L  N AE A ANAL  P12 3 1P11   MATAGGGGGGACATGGATTATAGGGGGGGGGGGGGGGAAAAGAAAAAAGGGGTTTCAGGTTTCAGGTGAAGGATATTTACAGTTGAACC  TITATCACCAGGGTGTACCTAMAAAGGGGGGGGGGGGGGGGACAAGAAAAAAAGGGGGTTTCACGTTTTTTTCCCCAAAGCCTCACCTTGATAAATGCAACCTTGAACC  TITATCACCAGGGTGTACCTAMAAAGGGGGGTGGGGGGGGGAAAAAAAAAA	
HAND BM M MNS PMA HA  INS MAP A SAPP FLEE FOR SEA BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER	Pag
HAND BM M MNS PMA HA  INS MAP A SAPP FLEE FOR SEA BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER BE BLE VER	
I HA L HINN NAE A AMAL D 12 3 1911  MATAGEGECEGACATGGATTITACGACGGCACCCCCCAGACGAGGAMAGAMAACGGCTTTCAGCTTTCGCAGTGAACATATTTACACTTGAACC  ITTATCCCCGGCTGTACCTAMAAGCTGCCGTGCGCGCTCCTCCTCTTTTTTTTTT	
NAE A ANAL 1911 1 1//  MATAGEGECGACATGGACTITTACGACGGECACGCGACGCAGAGGAAAAAAAAAAAAAAA	
MATAGGGCGAATGGATTITTAGGAGGGCAGGGCGAGGAAGGAAGAAGAAGAAGAAGAAGA	
MATAGEGETGACATGGATTTTTACGACEGGCACGGCAGGCAGAAAAAAAAAA	
THATCECGGCTGTACCTAMAAGCTGCCGTCTGCCCTTTTCTTTTTTCCCGAAAGTCCAAAGCCTCACTTGTATAAAGTGAACTTGAACATTGAAAAAGCTGCCGCTCTGCCCTTTTCTTTTTTTCCCGAAAGTCCAAAGCCCTCCAAAGTCCAAAAAGCTCCTGCGGAAAGCCTTAAAAAGCTCCTAAAAAAACGGGAATAAAACGGGAAAAAACGGGAAAAAACGGGAAAAAACGGGAAAAAA	
K. rhgflrrhaqtrkrkntkrkngfqlse.tylhlkp nsadmdfydgtrrrgkektafsfrsehiyt.n nsadmdfydgtrrrgkektafsfrsehiyt.n iaptwifttaradeekkkrlsafgvnifftlet lyrrcpnkrrcacvlflffpk.seshvykckf iaptwifttaradeekkkrlsafgvnifftlet lyrrcpnkrrcacvlflffpk.seshvykckf isgvhikvvarasssffrseaklkrlsci.vqf iflasmsk.apvrlrpfsfvaklkrlsci.vqf iagvhikvvarasssfffrseakptfmnvssv  HHB XMDBM M HNS PNA HA  1 NS HBP 1A B BLF VSL 1L  NAM CON NE O OAA UPU NU P11 211 12 23N 221 31  //  // // // // // // // // // // //	190
is a ptwifft are deekkkrls afgvniftlet  lyrrcpnkrrcacvlflfffpk.seshvykckf  lyrrcpnkrrcacvlflfffpk.seshvykckf  lflasmsk.spvrlrpfsfvaklkrlsci.vqf  .isgvhikvvarasssfffrseakptfmnvssv  HHB XMD BH M HHS PHA HA  IHS HBP IA B BLF VSL IL  P11 211 11 2 23N 2B1 31  //  //  //  //  //  //  //  //  //	
LYFTCPNKFTCacvlflffpk.seshvykckf  If lasmsk.spvrlrpfsfvaklkrlsci.vqf  .iagvhikvvarasssfffrseakptfmnvssv  HHB XHDBN M MNS PNA HA  1 HS HBP IA B B LF VSL 1 L  N AM CON ME D DA A UPU N U  P 11 211 11 2 2 3 N 281 3 1  //  AGGGGATTCAAATTCCTGATGTTTTGAAGATCCCTGTCTCTCCGTAAGTCCGTGATGTGGAAGGCGTTCAAGTGCCTTCGGAACATTCAAGTTCCAGAAGTCCCTTGTCGTAAGTTCAAGTTCAAGTTCCACAAAGTCCTTCGGAAACATCAAGTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTTCCACAAAGTCCCTTCGGAACAC  safkflmf.rs.rkksscrcmwkksvqlfteals  shsns.cfedpsgrrhhvgvcgrrfscsrchaagtcacataactcatagtcacataatacacttcaagtcacataacacataacacataacacataacacataacacataacacataacacataacacataacacataacacataacacacataacacacataacacacacacacacacacacacacacacacacacacaca	
HHB XMD BM M HNS PMA HA  I HS HBP 1A B BLF VSL 1 L  NAM OON HE O O AA UPU N U  P11 211 11 2 23 N 281 31  //  AGGGCATTCAAATICCTGATGTTTTGAAGATCCTAGGGGAAGGCATCATGTCGGTGATGTGGGAAGGCTTCAGCTGTTCACAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGGCTTCAGGAAGAACAC  S a f k f l m f . r s . r k b s c r c m w k a v q l f t e a l s  a h s n s . c f e d p s g r r h h v g v c g r r f s c s r k l c c  r i q i p d v l k i l a e e g i m s v y v e g g s a v h g s f v  l a n l n r i n q l d . r f f a d h r h i h f a t . s n v s a k  a c e f e q h k s s g l p l c . t p t h p l r n l q e r f s c  s r m . i g s t k f i r a s s p m m d t y t s p p e a t . p l k t  M M SNN AM N X H  B M PSL L S P M P  O	
HHB XMD BM M M NS PNA HA  I HS HBP IA B B LF VSL I L  N AM CON NE O DA A UPU N U  P11 211 11 2 2 3 N 2B1 3 1  //  //  //  //  //  //  //  //  //	
H HB XMD B H N	
I HS HBP IA B B LF VSL I L NAM OON NE O OAA UPU N U 211 11 2 23 N 2B1 31  // // GGGGATTCAMATICCTGATGTTTTGAAGATCCTTAGGGGGAAGAGAGGCATCATGTGGGAAGGCGGTTCAGCTGGAAGCTTTGTG  S a f k f l m f . r a . r k k a s c r c m w k a v q l f t e a l s a h s n s . c f e d p s g r r h h v g v c g r r f s c s r k l c c r i q i p d v l k i l a e e g i m s v y v e g g s a v h g s f v  l a n l n r i n q l d . r f f a d h r h i h f a t . s n v s a k c a c e f e q h k s s g l p l l e . t p t h p l r n l q e r f s q s r m . i g s t k f i r a s s p m m d t y t s p p e a t . p l k t  M SNN AM H X H B N PSL L S P H P O L HPA U E A N H  GGAAGGCTGTTTTCAGAAATCATCTTCTATTTTGCCCCTAAACTAATCGGAGGGAAGGCTGGAATTAGAGGCCACTTCCAAAGGTTTCAATCACCTTCTTTTTCCGACAAAGGTTTTCAATCACGGGGAAGGCTTTTCAAAGGTTTCAATCACCACAAAGGTTTTTAGTAGAAGAATCATCTTCTATTTTGCCCCTAAACTAATCGGAGGGAACGCAATGTCCCCAAGGTTTAATCTCCGGTGAAAGGTTTTCAAAAGTTCATCT  CK a v f k k s s s i l p l n . s e e r m l p a . s p v k v f n q e g c f q e i i f y f a p k l i g g t h a p s l i s g e g f q s	
NAM OON NE 0 DAA DPU NU NE 0 DAA 231 11 2 23 N 281 31 7/ // AGGGCATTCAAAATTCCTGAAGAAGAGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA	
P11 211 11 2 23 N 281 31 //  // /  /	
AGGGGATTCAMATTCCTGATGTTTTGAGGATCGCGGAAGAAGGCATCATGTCGGTGAAGGCGGTTCAGCTGTTCACGGAAGCTTTGTCGGGAAGGTTTAAGGATCGTAGGATCGCTTCTCCGGTAAGTTCACGGAAGCCTTTGTCGGGAAGGTTTAAGGGCGAAGCTTTGTCGGGAAGGTTTAAGGGCGAAGCTTTGTCGGGAAGGTTTAAGGGGGAAGCTTTGTCGGGAAGGTTTAAGGGGGAAGCTTTGAAAAGTTCGGAAGGTTTTCAGGAAGGCTTTTTGCCCGTAAGTTAAGCGGGGAAGCTTTAATCTCCGGTGAAGGTTTTCAATCAGGGGGAACGCAAGTTCCCAAAAGTTTCAATCAGGGGGAACGCAAGGTCGAATTAGAGGCCACTTCCAAAAGTTTCATCTCCGGAAGAAGTTTTTAATCAGGGGGAACGCAAGGTCGAATTAGAGGCCACTTCCAAAAGTTAGTAGCCCCCTAAACGTTAAGCGGGGAAGGTTTTCAATCAGGGGGAACGCAAGGTCGAAATTAGAGGCCACTTCCAAAAGTTAGGTCCCGAAAAGTTTTTAGCAAAAGTTTTTAGCCCCCTAAACTTAACCCTCCTTGCGTACGAGGGTCGAATTAGAGGCCACTTCCAAAAGTTAGTAGCCCACGAAAAGTTCTTTAGGAAGAGAAAAAAGTTTCAAAAAGTTAGAGGCCACTTCCAAAAGTTAGGTCCCGAAAAGTTTTAGAAGAGAAAAAAGTTAGAAGGCCACTTCCAAAAGTTAGGTCCCGAAAAGTTTAGAAGGCCACTTCCAAAAGTTAGGTCCCGAAAAGTTAGAAGGCCACTTCCAAAAGTTAGGTCCCGAAAAGTTAGAAGGCCACTTCCAAAAGTTAGGTCCCGAAAAAGTTAGAAGAAAAAAAA	
ICCCGTAAGTTTAAGGACTACAAAACTTCTAGGACCCCTTCTTCCGTAGTACAGCCGTTCAAGTGCCTTCAAAACTTCTAGAACCCTTCTCCGTAGTACAACCCTTCCGCCAAGTGCCTTCAAAACCTTCGAAACACCTTCCGCCAAGTGCCTTCAAAACACCTTCCGCGTAAGTTTAAGGACCAAGTGCCTTCGAAACACCTTCGCGTAAGTTTAAGGACCAAGTGCCTTCGAAACACCTTCGCGTAAGTTTAAGGACCAAGTGCCTTCGAAACACCTTCGCGTAAGTTTAAGGACCAAGTGCCTTCGAAACACCTTCGCGTAAGTTTAAGACCTTCTTTTTAAGAAACTTCTTTATTTTTGCCCCTTAAACTAATCGGAGGGAACCCAATGCCCCCAAACTTAATCCCGGTGAAAAGTTCTTCAAAAAGTTCTTTAGTAGAAACCGGGGGATTTGATTAGCCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTTCTAAAAAGTTCTTTAGTAGAAAACGGGGGATTTGATTAGCCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTTCAAAAAGTTCTTTAGTAGAAAACGGGGGATTTGATTAGCCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTTCAAAAAGTTCTTTAGTAGAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTCTTCAAAAAGTTCTTCAAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTTCAAAAAGTTCTTCAAAAAGTTCTTCAAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTCCTCCAAAAGTTCTTCAAAAAGTTCTTCAAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTCCAAAAGTTCTTCAAAAAGTTCTTCAAAAAGTTCTTCAAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTTCAAAAAGTTCTTCAAAAAGTTCTTCAAAAACGGGGGAATTTGATTAGCCTCCTTGCGTACGAGGGTCGAAATTAGAGGCCCACTTCCAAAAGTTCCAAAAGTTCCAAAAAGTTCTTCAAAAAACGGGGGAATTTGATTAGCCCTCCTTGCGTACGAGGGTCGAAATTAGAGGGCCACTTCCAAAAAGTTCCAAAAAGTTCCTCCAAAAAGTTCTTCAAAAAACGGGGGAATTTGATTAGCCCTCCTTGCGTACGAGGGTCGAAATTAGAGGGCCACTTCCAAAAAGTTCCAAAAAGTTCCTTCAAAAAGTTCAAAAAAAA	
laning in qld.rffadhrhihfat.snvsako acefeqhkssgiplie-tpthplrniqerfsq srm.igstkfirasspmadtytsppeat.plkt  M	q
B N PSL LS P N P  O L HPA UE A N H  2 1 1H3 11 2 1 1  //  AGAAGGCTGTTTTCAAGAAATCATCTTCTATTTTGCCCCTAAACTAATCGGAGGAACGCATGCTCCCAGCTTAATCTCCGGTGAAGGTTTTCAATCA  TCTTCCGACAAAAGTTCTTTAGTAGAAGATAAAACGGGGAATTGATTAGCCTCCTTGCGTACGAGGTCGAATTAGAGGCCACTTCCAAAAGTTAGT  C k a v f k k s s s i l p l n . s e e r m l p a . s p v k v f n q  r r l f s r n h l l f c p . t n r r n a c s q l n l r . r f s i /  e g c f q e i i f y f a p k l i g g t h a p s l i s g e g f q s	
B  N  PSL LS P H P  N  N  N  N  N  N  N  N  N  N  N  N  N	
D L HPA UE A N H 2 1 1H3 11 2 1 1  WAGAAGGCTGTTTTCAAGAAATCATCTTCTATTTTGCCCCTAAACTAATCGGAGGAACGCATGCTCCCAGCTTAATCTCCGGTGAAGGTTTTCAATCA  CTCTTCCGACAAAAGTTCTTTAGTAGAAGATAAAACGGGGAATTGATTAGCCCTCCTTGCGTACGAGGGTCGAATTAGAGGCCACTTCCAAAAGTTAGT  C k a v f k k s s s i l p l n . s e e r m l p a . s p v k v f n q  r r l f s r n h l l f e p . t n r n a e s q l n l r . r f s i f e g c f q e i i f y f s p k l i g g t h a p s l i s g e g f q s	
1 1H3 1 1 2 1  MGAAGGCTGTTTTCAMGAAATCATCTTCTATTTTGCCCCTAAACTAATCGGAGGAACGCATGCTCCCAGCTTAATCTCCGGTGAAGGTTTTCAATCA  CTCTTCCGACAAAAGTTCTTTAGTAGAAGATAAAACGGGGAATTGATTAGCCTCCTTGCGTACGAGGGTCGAATTAGAGGCCACTTCCAAAAGTTAGT  C k a v f k k s s s i l p l n . s e e r m l p a . s p v k v f n q  r r l f s r n h l l f c p . t n r r n a c s q l n l r . r f s i f  e g c f q e i i f y f s p k l i g g t h a p s l i s g e g f q s	
AGAAGGCTGTTTTCAAGAAATCATCTTCTATTTTGCCCCTAAACTAATCGGAGGAGGACGCATGCTCCCAGCTTAATCTCCGGAGGAGGTTTTCAATAACTAGAGGCCACTTTCCAAAAGTTAGTT	
TECTTCCCACAAAAGTTCTTTAGTAGAAGATAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAATTAGAGGCCACTTCCAAAAGTTAGT  ckavfkksssilpln.seermlpa.spvkvfnq rrlfsrnhllfep.tnrnacsqlnlr.rfsi; egcfqeiifyfapkliggthapslisgegfqs	
TECTTCCCACAAAAGTTCTTTAGTAGAAGATAAAACGGGGGATTTGATTAGCCTCCTTGCGTACGAGGGTCGAATTAGAGGCCACTTCCAAAAGTTAGT  ckavfkksssilpln.seermlpa.spvkvfnq rrlfsrnhllfep.tnrnacsqlnlr.rfsi; egcfqeiifyfapkliggthapslisgegfqs	¬
TETTCCCACAAAGTTCTTTAGTAGAAGATAAAACGGGGGATTTGATTAGCLLCLTTGGGAAAAGTTCTTTAGTAGAAGATAAAACGGGGGATTTGATTAGCLLCLTTGGGAAAAACGGGGGATTTGATTAGCLLCLTTGGGAAAAACGGGGGATTTGATTAGCLLCLTTGGGAAAAACGGGGGATTTGATTAGCLLCLTTGGGAAAAAACGGGGGATTTGATTAGCLLCLTTGGGAAAAAACGGGGGATTTGATTAGCLLCLTTGGAAAAAAAAAA	
ckav f k k s s i l p l n . s e e r m l p a . s p v k v f n q r r l f s r n h l'l f c p . t n r r n a c s q l n l r . r f s i ; e g c f q e i i f y f a p k l i g g t h a p s l i s g e g f q s	
erf from tifep to read the restriction of the restr	
erf from tifep to read the restriction of the restr	0
egcfqeiifyfapkliggtnaph	
,	
	-
lisnelf.rnqg.virifaheusirrhinei	
lspqk.simk.kaglsippvcaglkiepspk.d	

FIGURE 3G

		Page
F 1 N 1	////	BM FN SA OL TE KA F 5 TRANSLATION E3 13 STOP RBS , ORF 4
TGAAAGATGTCCCCTTATTACAATTCACTGATATAACCCAA		
ACTITICTACAGGGGAATAATGTTAAGTGACTATATTGGGTTT		
.kmspyynsli.pk ercplitih.ynpn kdvpllqftditqi	rp.yqtdgkt igrdikltækj	dkgigw.pe otke.dgdhy
h fidglesiygl slhgrivi.qylgi ifstgkncnvsivw	datid for cf frghy, vaply	r clflit v m / slpiphhgh
H T H THH N P A B ANN L H Q O QFF A 1 1 2 131 3	SNN HH N PSL AA S HPA EE E 1H3 13 1	M M M H S . N A A E L M M H
TTACAGGAATTATCGAAGAAACAGGCACAATCGAATCCATGA ATGTCCTTAATAGCTTCTTTGTCCGTGTTAGCTTAGGTACT	TTTTTCGTCCCGTACGTTACCGGAATTGAT	ZATTTACGAGTTTCTAAAATCTCCTACA
tgiiee-tgtiesmilqeiskkqaqsnp. yrnyrrnrhnrihe	kkqgmqup.lksracnglny	lnaqrf.rmf .mlkdfrgc
nvpiissvpvisda kesndffcaedfgh .lf.rlfletriws	ffapcaiak v ffcpmchg, s	îlhe fîksst nfa.ln.li
F O K 1	H A E 3	F O K 1
TCATCTTGGCGACAGCATTGCAGTGAACGGCATTTGTCTGAC		2400
h l g d s i a v n g i c l t i l a t a l q . t a f v . l s s w r q h c s e r h l s d . r p s l m a t f p m q r v	slilqkinso ch.fykksih	wmlckqsk sgcya.nsqs
	s d s i k c f i t e	chinhrfedf

FIGURE 3H

MH	H	•	F	XGGCHM	нн	D	B HH	NS
AA .	S		N	MDDFAN	<b>AA</b>	D	SAA	LA
E	E		υ	AIIREL	EÉ	E	H EE	AL
3	1		н	322131	13	1	2 13	31
				[]]]]	1		1	1
		AAGTAAATCTGGAAAGAGC						
	•	TTCATTTAGACCTTTCTCG						
		v n l e r a						
		k.i w ker sksgks					q <b>a</b> :	
	-		-					
		ltfrsla						
		fyiqfs						
. tvsh	nilite	lldpfl	<b>s</b> p p	ırgn	rign	rr	l g	n
	H KE	SX	Ю	AAN	M E		HD	
	N SA	PB	SR	VSL	S C		SR	
	F PR	<b>00</b>	EA	ALIA	E O		EA	
					1 D		11	
	3 21	12	_ 11		. •			
ACGGAACTGCGGA	AATCACACGAATTGAAGA	P2 PROHOTI	TATGATTTAA	AAATGGACCCGT	CATTAACAAA			
ACGGAACTGCGGA TGCCTTGACGCCT	ATCACACGAATTGAAGAC TTAGTGTGCTTAACTTCTC	PROHOTI GAMAGCAACGCAGTTTAC CTTTTCGTTGCGTCAAATG  k s n a v y k a t q f t	ER ETATGATTTAA FATACTAAATT y d l k m i .	MATGGACCCGT	CATTACAAA GTAATTGTTTT	TTGTAA	CCAAAA	: III
TGCCTTGACGCCT	ATCACACGAATTGAAGAG TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r	PROMOTI GAMAGCAACGCAGTTTAC CTTTTCGTTGCGTCAAATG	ER ETATGATTTAA EATACTAAATT  y d l k m î f k	AAATGGACCCGT PTTACCTGGGCA  m d p s k w t r n g p v	CATTACAAA GIAATIGITII . l t k h . q k i n k i	TTGTAA	CCAAAA V ( g f	: TTT k r k
ACGGAACTGCGGA TGCCTTGACGCCT  g t a e t e i r k r n c g i	ATCACACGAATTGAAGAG TTAGTGTGCTTAACTTCTC  i t r i e e s h e l k r n h t n . r i	F22 PROHOTI GAAAAGCAACGCAGTTTAC CTTTTCGTTGCGTCCAAATG  k s n a v y k a t q f t e k q r s l l f l l a t	ER ETATOATTIAA  MATACTAAATT  y d l k  m i f k  . s k	AAATGGACCCGT	GTAATTGTTTT	t i h w	CCAAAA V [ g f	k r k k
gtae etrk rncgi	ATCACACGAATTGAAGAI TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r  n h t n . r i  i v r i s s f d c s n f	F I Lat.	ER  TATGATTAAA  AATACTAAAATT  y d l k m i f k f k	AAATGGACCCGT TTTACCTGGGCA  m d p s k w t r n g p v	GTATTATCATA	t l h w	CCAAAA v l e f . g f t k q n	k r k
gtaeelrk rncg:	ATCACACGAATTGAAGAI TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r  n h t n . r i  i v r i s s f d c s n f	F22 PROHOTI GAAAAGCAACGCAGTTTAC CTTTTCGTTGCGTCCAAATG  k s n a v y k a t q f t e k q r s l l f l l a t	ER  TATGATTAAA  AATACTAAAATT  y d l k m i f k f k	AAATGGACCCGT TTTACCTGGGCA  m d p s k w t r n g p v	GTATTATCATAL  GTAATTGTTT  Ltk inki inki dnvf	t l h w	CCAAAA v l e f . g f t k q n	k r k
gtaeeirk rncg:	ATCACACGAATTGAAGAG TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i  f . v f q l	FROMOTI  GAAAAGCAACGCAGTTTAG  k s n a v y k a t q f t k q r s i l f l l a t . i f a v c n v s f c r l k  HH 8	ER CATTAGATTAAA  CATACTAAAATT  y d l k  m i .  f k  f k  s k  / i i .  s h n l	AAATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t	GTTAACAAA  GTAATTGTTT  Ltk h.qk inki dnvf	ti hw ni vn fc	V ( g f . g f . g n p k	k
GCCTTGACCCCT  gtae elrk rncgi	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i	FARAMAGCAAGGAAGTTTAG  K s n a v y k a t q f t k q r s i i f i l a t . i f a v c n v s f c r l k	EXECUTATION TO THE PROPERTY OF	AAATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t	GTATTATCATAT  Ltk h.qk inki dnvfci	ti hw ni vn fc fm	v l g f . g f	k
gtae etrk rncgi pvas vssr	ATCACACGAATTGAAGAG TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i  f . v f q l	F2 PROHOTI GAAAAGCAACGCAGTTTAC  k s n a v y k a t q f t k q r s l l f l l a t . l f a v c n v s f c r l k  HH    B NA    B FE  v	ER ETATGATTTAA  BATACTAAATT  y d l k m i . f k . i i . s h n l  H HM A BB E CO	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D K P S	GTTAACAAA  GTAATTGTTT  Ltk h.qk inki dnvf	ti hw ni vn fc fm	V ( g f . g f . g n p k	k
gtae elrk rncg pvas vssr	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r  n h t n . r i  i v r i s s f d c s n f i f . v f q l  A FM L OS U KE 1 11	F C C C K  HH B  NA B	ER ETATGATTTAA  ATACTAAATT  y d l k m i f k f k s h s h n l  H MM A BB	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D M P S N E	GTATTATCATAT  Ltk h.qk inki dnvfci	TTGTAA	CCAAAA  V (  g f  t k q n p k	# F K F K F K F K F K F K F K F K F K F
gtae etrk rncgi pvas vssr	ATCACACGAATTGAAGAI  IT T i e e s h e l k r n h t n . r i i v r i s s f d c s n f l f . v f q l  A FM L OS U KE 1 11	F2 PROHOTI GAAAAGCAACGCAGTTTAC  k s n a v y k a t q f t k q r s l l f l l a t . l f a v c n v s f c r l k  HH    B NA    B FE  v	ER  CTATGATTTAA  ATACTAAATT  y d l k  m i . f k . s k / i i . s h n l  H HH A BB E 00 3 21	AAATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S N E 1 1	GTATTGTTT	ti hwni vn fc fm	CCAAAA  V {     f .     g f .     q n     p k     N       8    0    1	TTT k r k f l l l l l l l l l l l l l l l l l l
GCCTTGACCCCT  g t a e e l r k r n c g i p v a s v s s r r f q p	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i  f . v f q i  A FH  L OS  U KE  1 11  GATGGCGTGAGCTTAACCJ	FE PROHOTT  GAAAAGCAACGCAGTTTAC  k s n a v y k a t q f t k q r s l l f l l a t . l f a v c n v s f c r l k  HH	EXTACATTAAATT  y d l k m i f k . s k , i i s h n l  H HH A BB E CO 3 21	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f 1 s g f h v r f p g t  D H P S H E 1 1	CATTAACAAA  GTAATTGTTT  L t k h q k i n k i d n v f	t i h w m i i v n i f c f m	CCAAAA  v l f . g f p k  N ii B O1	k r k f
ACGGAACTGCGGA  GCCTTGACCCCTT  g t a e t e l r k r n c g i s p v a s v s s r r f q p	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i  f . v f q i  A FH  L OS  U KE  1 11  GATGGCGTGAGCTTAACCJ	FE V  ATATTCGGCCTGACAGAAG  ATATTCGGCCTGACAGAAG  PROMOTT  A	EXTACATTAAATT  y d l k m i f k . s k , i i s h n l  H HH A BB E CO 3 21	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f 1 s g f h v r f p g t  D H P S H E 1 1	CATTAACAAA  GTAATTGTTT  L t k h q k i n k i d n v f	t i h w m i i v n i f c f m	CCAAAA  v l f . g f p k  N ii B O1	k r k f
ACGGAACTGCGGAA TGCCTTGACGCCT	ATCACACGAATTGAAGA  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r  n h t n . r i  v r i s s  f d c s n f i  f . v f q l  A FN  L OS  U KE  1 11  GATGGCGTGAGCTTAACCC  CTACCGCACTCGAATTGGT	FE V  ATATTCGGCCTGACAGAAG  ATATAAGCCGGACTGTCTTC	ER ETATGATTTAA  ATACTAAATT  y d l k  m i f k  f k  f k  s h n l  H HN  A BB  E OO  3 21	MAATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S N E 1 1  GATCTCCTTAAT	CATTAACAAA  GTAATTGTTT  Ltk h.qk inki dnvfc: mLL  M( Bi O) 1:	t t t h w n i i v n n	GAAAACG	k r k r l l l l l l l l l l l l l l l l
ACGGAACTGCGGAA  TGCCTTGACGCCTT	ATCACACGAATTGAAGA  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r  n h t n r r  i v r i s s  f d c s n f l  A FM  L OS  U KE  1 11  GATGGCGTGAGCTGAATTGGG  d g v s l t	FE V  TATAAGCCGGGCTGTCTTC  if g l t e c	ER ETATGATTAAATT  y d l k m i f k f k s h n l  H HM A BB E OO 3 21  SACACAGTGAC  ETGTGTCACTG	AAATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S H E 1 1  GATCTCCTTAAT  CTAGAGGAATTA	GTATTATCAAAA  L t k h . q k i n k r d n v f c : m   l  ACCGCATACG	t t t h w n i i i i i i i i i i i i i i i i i i	GAAAACG	k f f D D P N N 1 1 TAG
ACGGAACTGCGGAACGCCTT  g t a e t e t e t e t e t e t e t e t e t	ATCACACGAATTGAAGAI  IT T i e e s h e l k r n h t n . r i v r i s s f d c s n f i f . v f q l  A FH L OS U KE 1 11  GATGGCGTGAGCTTAACCJ  CTACCGCACTCGAATTGGG d g v s l t m e . e . p	FE V  TATAMAGCAGGACTGCCTCC  ATTATCGGCCTGACAGAGG  TATAAGCCGGACTGCCTCC  F G L C C C C C C C C C C C C C C C C C C	ER  CTATGATTTAA  ATACTAAATT  y d l k  m i f k  f k  f k  s h n l  H HM  A BB  E CO  3 21  SACACAGTGAC  CTGTGTCACTG	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S H E 1 1  GATCTCCTTAAT  CTACAGGAATTA  i s l i s p y	GTAATTGTTTT  Ltk h.qk ink ink dnvfc mLL  ACCGCATACG	t t t h w n i ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '	CCAAAA  v l  f .  g f .  g n  p k  R  GAAACG	k r f l D D P N 1 1 ATC
ACGGAACTGCGGA  TGCCTTGACGCCT	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i f . v f q l  A FH L OS U KE 111 / GATGGCGTGAGCTTAACCC  CTACCGCACTCGAATTGGT d g v s l t m a . a . p	FE V  TATAAGCCGGACTGTCTTC  If g t e d  y s a . q k  i r p d r r	EXTACTATATA  SATACTAAATT  y d l k  m i  f k  . s h  H HM  A BB  E OO  3 21  CACACAGTGAC  CTGTGTCACTG	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S H E 1 1  GATCTCCTTAAT  CTACAGGAATTA  i s l i s p y d l l n	GTATTGTTTT  Ltk h.qk ink ink ink dnvfc mll  ACCGCATACG	t i h w n i f c f m	CCAAAA  of f  g f  q n  H  B  o  t  ccattricc  ccattricc  ccattricc  ccattricc  cattricc  cattri	k r k l D D P N 1 1 TAGC i s l
ACGGAACTGCGGAACTGCCTTGCCTTGACGCCTT  g t a e t e t r k r n c g i s p v a s v s s r r f q p  ATCAATTACTGTGG  ATCAATTACTGTGG  TAGTTAATGACACC	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r n h t n . r i  i v r i s s f d c s n f i  A FM L OS U KE 111  GATGGCGTGAGCTTAACCC  CTACCGCACTCGAATTGGT  d g v s l t m a . a . p	FRONOTT  GAAAAGCAACGCAGTTTAG  k s n a v y k a t q f t k q r s l l f l l a t . l f a v c n v s f c r l k  HH	EXTACATTAAATT  y d l k m i f k . s h n l  H HHH A BB E CO 3 21  CACACAGTGACC  CTGTGTCACTG	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S H E 1 1 GATCTCCTTAAT  CTACAGGAATTA  i s l i s p y d l l n	CATTAACAAA  GTAATTGTTT  L t k h . q k i n k ! d n v f c ! m L l  ACCGCATACG  TGGCGTATGC:	t i h w n i ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '	CCAAAA  V l  f .  g f  Q n  N  B  O 1  CCAAAA  CCTTTGC  e t  n d	k r k l l D D P N 1 1 ATC
g t a e t e t e t e t e t e t e t e t e t	ATCACACGAATTGAAGAG  TTAGTGTGCTTAACTTCTC  i t r i e e  s h e l k r  n h t n . r i  i v r i s s  f d c s n f i  v f q i  A FN  L OS  U KE  1 11  GATGGCGTGAGCTTAACCC  CTACCGCACTCGAATTGGC  d g v s l t  m e . a . p  w r e l n h	FE V  TATAAGCCGGACTGTCTTC  If g t e d  y s a . q k  i r p d r r	EXTRACTABATT  y d l k m i y d l k m i f k f k f k s h n l s h n l s h n l s t v t v t v v v v v	AMATGGACCCGT  TTTACCTGGGCA  m d p s k w t r n g p v  f i s g f h v r f p g t  D H P S N E 1 1  GATCTCCTTAAT  CTAGAGGAATTA  i s l i s p y d l l n	CATTAACAAA  GTAATTGTTT  L t k h . q k i n k i n k i n k i T t a y d  GTAATTGTTTT  GTAATTGTTTT  ACCGCATACGC  P h t t a y d	t l h w f c f m  ATCAGC  TAGTCG i s s a q i l	GAAAACG GAAACG CCTTTGC E K r d	k r k k T D P N N 1 T T T T T T T T T T T T T T T T T

FIGURE 31

					Page 10
H DP	TH B	H D	CT		
B PV	AN S		LA LA	ε C	
ONU	OF N		AQ	P	
1 11	13 1	1 1	11	1	
	,		,		
TTTTCAGAAAAACGATCGGCTCTAAAGTGAAT				CATAMAGECAATGAAAATAAGACEC	2800
AAAAGTETTTTTTGCTAGCCGAGATTTCACTTA					3533
fsektigskyn fqkkrsalk, i frkndrl, sey	s n a	i.se ( ydrk	ymyrfil nicidfo iyvsifa	ikpmkirp	
k, ffrdarfhi	dfa	siip ihds	•		
Ť AM	н	F	MHN	н мон	
T FS	N ORF	٥ ،	BSL	8 SR G	
H LE	TRANS	LATION K	QPA	O EA A	
2 21	1 STOP	1	2H3 ORF 3 //	2 11 1	
AACAAACCATTACAAAAGCCTTCTTAAGCGAAA	ACCCCTTT				
****, *************************					2900
TTGTTTGGTAATGTTTTCGGAAGAATTCGCTTT	•		ACGTACAAAGTAGGCTATCT	•	2700
qtitkaflsernkplqkps.aktnhyksllkrk	ngf. tafr rtt	rgrfeedterki erki erki erki erki	acfir.k hvssdri mfhpie ahkmry	khwtl.kkrstgrfkkrealdalkk	
8	H	B	SNN	H M	
8	В	S	PSL	P \$	
V	0	H	HPA	A E	
2	2	2	183	2 1	
GGCGAAGTCATCATCGTTGTAGATGATGAAGAC	AGAGAAAATG	AGGAGACTTTG	// GGCTCTTGCCGAGCATGCA		7000
CCGCTTCAGTAGTAGCAACATCTACTACTTCTC	TCTCTTTTAC	TTCCTCTGAMC			3000
aksssl.mmkt rshhrerre	e k m !		ellps mq sscracn	rrksttlu	
afddnyiifv lrlrqlhhl pstmmttsss	sfi clfh lsf	llsq	•	rrfdnykh lapt.s.p ygstmlki	

FIGURE 3J

						Page
. и и	M D	FH D	и нир		рн н	
N L	8 P	NN D	8 P B P		LPN	
L A	O N	ULE	о ном		E H F	
1 3	1 1	H1 1	2 1 1 1		111	
			AGGAAATCGCAGACAGGCTTGATCTT	+,		3100
GTGTACCETET	CCTGACTAGAC	GTGCGGCGAGTCAC	CETTTAGEGTETGTEEGAACTAGAA	GTGGGATACCAACTCGTA	FTATGTCTGAGAGTGGT	
t w e r	tdl	haaq. Inlse	rksqtglif gnrqa.ss eiadrldl	pyg.a. hpmveht	ntdshh	
c m p s		arret	lfdevpkik pfrleaqd ssiastssr	vrhnlm eg.pqa	icvrv yylseg	
		E B	и нени хвис	т т		
		С В	B ICHA HGBF			
		P V	O NAME OLO			
		1 2	2 P712 2211			
			/// GACAGGTATCAGCGCTCAAGAAAGA	PCTTTTACCGTTCAAGCAT	TGCTGGACAGCAAATC	c c
ciyr	khr: veid	ps.ne hretk	qvsalkkd dryqrsrki tgisaqer	tyrs. sftvqal	ildsks	•
s c k 9	y & y '		totdastfs sty.retf rvpita.st	rkgnlc ik.rels	qqva f	9
H	BHNS	SB N	H	KH HTASH	AF EHD	
G		ES S	K	SI HHSBA	LLN CBP	
A	AAIR		L	PN AAUVE	NUU 10 N	
1		1J 2	1	1P 11113 //	N1H 5 1 1	
GCCATCTGATI	TTCAGCGTCCC	'// IGGGCACATTTTCC	acteattecgaaaaaaggaggtgtcc	TGAAAAGCGCGGGCCATA	CAGAAGCTGCTGTTGAT	re
					<del></del>	33
ACGGTAGACTA	WAGTEGEAGGE	CCCGTGTAAAAAGG	TGACTAACGETTTTTTTCCTCCACAGC	ACTITICGCGCCLGGIAI	BICTICOACOACAACTA	
ai.f	5 4 5 5	ahfs	, irkkevs tdcekrrcp iiakkggv	ekrgpy	F S C C	•
	farp	ghifp				
hwri	k i t :	pynk nacke	wqnrffstd vsqsfllh: siaffppt	qfaram asfrpgy	cfs s n	i d

FIGURE 3K

					Page 12
					Page 12
H A BXMD EHAS		8		D A SKEN	
I L IHBP CGPC N U NOON RAYR		8		D L SGAS	
3 1 1211 2111		V 2		E U TINP 1 1 1A22	
1 11		•	•	11/	
TTGCTGAAGCTTGCGGATCTCCA	GGAGECGGCGTCATTTGTGA	AATTATGAATGAAGACGGA	ACGATGGCGAGAGT	GCCTGAGCTCATTGAAATTG	GAA
AACGACTTCGAACGCCTAGAGGT	CETCGGCEGEAGTAAACACT	TTAATACTTACTTCTGCCT	TGCTACCGCTCTCA	CGGACTCGAGTAACTTTAACC	3400 CTT
ilkladiq c.slrisr aeacgsp:	srrhl.; gagvice	nye rrn imaneda:	dges: tmarv	a.ah.nc pelieia	e k
ksfsasru	s a a d n t	fnhifv -	<del></del>	······································	
qqlkrie	llrr.kh	f.shlr.	fspsl	* a * . a f a	•
rasaqpdg	papt mqs	iifssp	vialt	gssmsia	ı f
H SH BHD	и н	M	T	B H F	
S FP CBP	S N	Ä		B I N	
E AH LON	E F	Ε		V N U	
1 11 111	1 1	3	1	1 2 H	
AAAGCATCAATTAAAAATGATCA	^^1T146C4TTTC17+^4*	710000710117870101.			
			CACIIGICGAGCGI	GAAGTTGACATTACGCTGCCT	ACT 3500
TTTCGTAGTTAATTTTTACTAGT	GTAATTCCTAAACTAAGTT	LTGGEAATGTTAGACTGTT	TGAACAGETEGEA	ETTCAACTGTAATGCGACGGA	+ 3300 TCL
sin.k.s <sub>i</sub>	olri.fn	tvti.ql	1 l s s v	kltlrel	ŧ
sin.k.s kasikndh khqlkmit	olri.fn h.gfdsi ikdlig	tvti.ql piqsdn /rvnitt	ilssv tors.	kltlrel s.hysay	
sin.k.s kasikndh khqlkmit	olri, fn h.gfdsi ikdliq	tvti.ql plqsdn yrynltt	tera.	kltlrcl s.hyaay evditlp	t *
sin.k.s kasikndh khqlkmit flml.fhd	olri.fn h.gfdsi ikdliq y	tvti.ql plqsdn yrynltt 	t c r a .  i v e r e	kitirei s.hyaay avditip	t
sin.k.s kasikndh khqlkmit flml.fhd fadilfs,	plri fn h.gfdsi ikdliq y gnliqnl w.pnsei	tvti.ql plqsdn yrynltt  vtviqc	tera. ivera. ckdit	kitirei s.hysay svditip fnvnrqr	t t
sin.k.s kasikndh khqlkmit flml.fhd	plri fn h.gfdsi ikdliq y gnliqnl w.pnsei	tvti.ql plqsdn yrynltt  vtviqc	tera. ivera. ckdit	kitirei s.hysay svditip fnvnrqr	t t
sin.k.s kasikndh khqlkmit flml.fhd fadilfs, fc.nfii	plri fn h.gfdsi ikdliq y gnliqnl w.pnsei ymlski	tvti.ql plqsdn yrynltt  vtviqc	tera. ivera. ckdit	kitirei s.hysay svditip fnvnrqr	t t
sin.k.sgkasikndhkhqlkmit flml.fhd fadilfs. fc.nfii	plri fn h.gfdsi ikdliq j gnliqnl w.pnsei mlski	tvti.qp plqsdn yrynltt  vtviqc igncdsl yr.lrvv	tera. ivera. ckdit	kitirei s.hysay svditip fnvnrqr niqe.aa stsavsg	t
sin.k.sgkasikndh kasikndh khqlkmit  flml.fhd fadilfs, fc.nfiiv	plri fn h gfdsi ikdliq gnliqnl W pnsei / mlski	tvti.qp plqsdn yrynltt vtviqe igncdsl yr.lrvv	tera. ivera. ckdit	kltirel s.hysay svditip fnynrqr niqe.aa stsaysg	t
sin.k.sgkasikndhkhqlkmit flml.fhd fadilfs. fc.nfii	plri fn h.gfdsi ikdliq j gnliqnl w.pnsei mlski	tvti.qh plqsdn yrynltt  vtviqc igncdsl yr.lrvv  NN SL PA	tera. ivera. ckdit	kltirel s.hysay vditip fnynrqr ilqe.aa stsaysg	t 
sin.k.sgkasikndhkhqlkmit flml.fhd fadilfs, fc.nfii	plri.fn h.gfdsi ikdliq gnliqnl w.pnsei miski.	tvti.qt piqsdn yrynitt vtviqc igncdsi yr.irvv	tera. tera. tvere ekdit vqrah	kltlrcl s.hyaay avditlp fnvnrqr nlqc.aa stsavsg HCH GFP ERA 222	t s
sin.k.sgkasikndhkhqlkmit flml.fhd fadilfs, fc.nfii  fM. IS NE 11	plri.fn h.gfdsi ikdliq y gnliqnl M.pnsei mlski.  M N L 1	tvti.qt piqsdn /rynltt	tera. lvera. ekdlt vqrał vstsr	k l t l r c l s . h y a a y s v d i t l p f n v n r q r n l q c . a a s t s m v s g HCH GFP ERA 222	t 
sin.k.s kasikndh khqlkmit flml.fhd fadilfs, fc.nfiiv FM iS NE 11	Plri fn h.gfdsi ikdliq y n liqnl N.pnsei mlski M N L 1	tvti.qi piqsdn yrynltt vtviqc igncdsl yr.lrvv  NN SL PA H3 /	tera. lvera. ekdlt vqraf vstsr	k it ire is a y a a y a v d it ip for vor q r on it q c a a s s t s m v s g HCH GFP ERA 222	t
sin.k.sgkasikndhkhqlkmit  flml.fhd fadilfs. fc.nfii  FM IS NE 11  GATTTTGGGACATTTAAGGTTTAT  CTAAAACCCTGTAAATTCCAAATA  ilghlrfm fwdi.glwdfgtfkvy	cctatgtgtttactccatcs  d t q m r . m  g y t n e v d	t v ti. q h piqsdn / r y n i t t  v t v i q c i g n c d s i y r . i r v v  NN SL PA H3 / ITGGAAAAGAGCATGTCGCA ACCTTTTCTCGTACAGCGT	t c r a .  l v e r e  c k d l t  v q r a f  v s t s r  AAACACTACCCTCT	k it irc is a yaaya ya ya it ip  f n y n r q r n iq c a a s s t s m y s g  HCH GFP ERA 222  ATGTGCCGTTCGGAGAAGAAC  ACACGGCAAGCCTCTTCTTG	t
sin.k.sgkasikndhkhqlkmit  flml.fhd fadilfs, fc.nfii  FM IS NE 11  GATTTTGGGACATTTAAGGTTTAT  CTAAAACCCTGTAAAATTCCAAAATA  ilghlrfm fwdi.glwdfgtfky	Deri fn h g f d s i i k d l i q n g n l i q n l M . p n s e i / m l s k i .  GGATACACAAATGAGGTAGA CCTATGTGTTTACTCCATCT d t q m r . n y i h k . g r g y t n e v d	t v t i . q i p i q s d n / r y n l t t  v t v i q c i g n c d s l y r . l r v v  NN SL PA H3 / ACCTITICTCGTACAGGGT i e k s m s h w k r a c r i g k e h v a	tera. lvera. ekdlt vqra. yqra. ystsr  AAACACTACCCTCT l.wer edgr fvmgd	k it ircl s.hyaay v ditip fn v n r q r ilqc.aa sts m v s g  HCH GFP ERA 222 ATGTGCCGTTCGGAGAAGAAC ACACGGCAAGCCTCTTCTTG icrsekn: cavrrrt v p f g e e p	t
sin.k.sgkasikndhkhqlkmit  flml.fhd fadilfs. fc.nfii  FM IS NE 11  GATTTTGGGACATTTAAGGTTTAT  CTAAAACCCTGTAAATTCCAAATA  ilghlrfm fwdi.glwdfgtfkvy	Deri fn h g f d s i i k d l i q y g n l i q n l W . p n s e i y m l s k i   GGATACACAAATGAGGTAGA  CCCTATGTGTTTACTCCATCT  d t q m r . m y i h k . g r g y t n e v d  s v c i l y h i c l h p l	tvti.qh plqsdn /rynltt vtviqc igncdsl yr.lrv  NN SL PA H3 / ITGGAAAAGAGCATGTCGCA ACCTTTTCTCGTACAGCGT	tera.  tera.  tera.  tvera.  ckdlt  vqra.  vqra.  ATTTGTGATGGGAGA  AAACACTACCCTCT  lwer  cdgr  fvmgd	k it ire is a y a a y a v d it ip  fn vn r q r il q c a a st s m v s g  HCH GFP ERA 222  ATGTGCCGTTCGGAGAAGAAC  ACACGGCAAGCCTCTTCTTG C r s e k n c a v r r r t v p f g e e p	ccc 3600 ccc

FIGURE 3L

AA HNS B	B H HA	н	В	RAA	F PNAF	FBTKHH	. F	
VS PCC S	S A AF	e P	8		N VSLN		N	
AU AIR H	N E EL	H	٧	RAU	U UPUU	UVAPAN	Ü	
21 211 1	2 3 23	1	1	221	H 281H	41111P	Я	
/ /	. /			//	///	/		
	GAATGTCTCACAGGTGACGT							700
	CTTACAGAGTGTCCACTGCA							
	nvsgvto					. p r . t	k i p	
	mshr.r					•	n c	
	ecltgdv 							
	ftectv							
	lid. Lhr						g f a	
ntrtce	s h r v p s t	. n p e	r q <b>s</b> q	P 9	c s c	f	wia	
	•							
HDDSB8 I	HR R H H M		MS	×	н 🛦	на		
	GS S I H N		SF			IL		
	IP A HAL		EA	_		N U		
3111JJ	A2 1 P1 1		1N	1	31	3 1		
			176					
	/							
AGAAGGCCGTGGAGTGC	TCCTGTACTTGCGCCAAGAA		TTAATCAAT					
AGAAGGCCGTGGAGTGC	TECTGTACTTGCGCCAAGAA	,	TTANTCANTA			·	3	800
AGAAGGCCGTGGAGTGC	TCCTGTACTTGCGCCAAGAA	,	TTANTCANTA			·	3	800
AGAAGGEEGTGGAGTGE	TECTGTACTTGCGCCAAGAA	CCTGCTCCGTAGCC	TTTAATCAATA	TTAATT	TTCGAATA	ATTCGAAGTCCTTG	TTCCGATA	800
AGAAGGEEGTGGAGTGE	TECTGTACTTGEGECAAGAA AGGACATGAACGEGGTTETT	CCTGCTCCGTAGCC	TTTANTCANTA	 ITTAATT	TTCGAATA	STEGAAGTCCTTG	TTCCGATA	800
TOTTCCGGCACCTCACG	AGGACATGAACGCGGTTCTT  C C C C A K K  P V L A P P F  L y L P Q C	deas v	TTTMTCATA	TTAATT	TTCGAATA	sfrn asgt klqeq	kam Pl.	800
AGAAGGCCGTGGAGTGC	AGGACATGAACGCGGTTCTT  C C C C C C C C C C C C C C C C C	d e a s v	TTTANTCAATA	n . i k	TTCGAATA	S f r n . a s g t k l q e q	k a m	8800
AGAAGGCCTGGAGTGC	AGGACATGAACGCGGTTCTT  C C C C A K K  P V L A P F F  L y L F q e	deas v trhr grgig	TTTMTCMT/ 	n . i k c i k	TTCGAATA k l i s l . a y	S f r n . a s g t k l q e q	k a m r l . i g y	800
AGAAGGCCGTGGAGTGC  TCTTCCGGCACCTCACG  k a v e c : r r p w s a e g r g v l  c f a t s h l l g h i a	AGGACATGAACGCGGTTCTT  S C t C a k k  p v l a p r r  l y l r q e  e q v q a l f  g t s a g l	deas v trhr grgig	MATTAGTTA:  s i f n q . l i n i	TTAATT  n . i k c l k c l k	TTCGAATA k l i s l , a y f s i l k	sfrn.asgt klqeq lklf	k a m r l . 19 y	800
AGAAGGCCGTGGAGTGC  Kayec: rrpwsa egrgvl  cfatsh llghia	AGGACATGAACGCGGTTCTT  C C C C A K K  P V L A P F F  L y L F q e	deas v trhr grgig	MATTAGTTA:  s i f n q . l i n i	TTAATT  n . i k c l k c l k	TTCGAATA k l i s l , a y f s i l k	sfrn.asgt klqeq lklf	k a m r l . 19 y	8800
AGAAGGCCGTGGAGTGC  TCTTCCGGCACCTCACG  k a v e c : r r p w s a e g r g v l  c f a t s h i l g h i a s p r p t :	AGGACATGAACGCGGTTCTT  S C t C a k k  p v l a p r r  l y l r q e  e q v q a l f  g t s a g l	deas vot nhrige ngig	TTTATCATA	TTAATT  n . i k c l k c l k	TTCGAATA  k l i s l s l f s i l k f a	sfrn.asgt klqeq lklf	k a m r l . 19 y	5800
AGAAGGCCGTGGAGTGC  K a v e c : r r p w s a e g r g v l  c f a t s h l l g h l a s p r p t :	AGGACATGAACGCGGTTCTT  S C t C a k k  P V l a P r r  l y l r q e  • q v q a l f  g t s a g l  s r y k r w s	deas vot nhrige ngig	MATTAGTTA:  . s i f n q . l i n i t . d i n l . y	ittaatt nik clk 	TTCGAATA  k l i s l , a y  f s i l k f a ,	sfrn.asgt klqeq tlklf yaepv	k a m r l	3800
AGAAGGCCGTGGAGTGC  TCTTCCGGCACCTCACG  k a v e c : f r p w s a e g r g v l  c f a t s h l l g h i a s p r p t :  H B N G	AGGACATGACGGGGTTCTT  SCCCA & k  PVLapr  Lylrqe  eqvqalf  gtsagl  srykrws  HHH H M HBXM  1 HA N B PINB  N AE F O ANOC	deas ver to her services and service representations and service representations and services representations are services representations and services representations and services representations and services representations and services representations and services representations and services representations are serviced representations and services representations are services representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services representations are serviced representations and services represent	MATTAGTTAT  . s i f n q . l i n i . t . d i n l . y k i l	i k c l k c l k f . f . f i t l n	TTCGAATA  k l i s l . a y  f s i l k f a .	S f r n a s g t k l q e q t l k l f y a e p v l s - s	k a m r l . I g y	5800
AGAAGGCCTGGAGTGC  K a V e C : r r p w s a e g r g v l  c f a t s h l l g h l a s p r p t :  H B N G L	AGGACATGACGGGGTTCTT  SCCCA & k  PVlapr lylrqe  eqvqalf gtsagl srykrws  HHH H M HBXM 1 HA N B PIMB NAE F D ANGO P12 1 2 2121	deas v t r h r g r g i g s s a d i l v l c r p r p m p	MATTAGTTAT  . s i f n q . l i n i . t . d i n l . y k i l	i k c l k c l k f . f . f . f . f . f . f . f . f . f .	TTCGAATA  k l i s l . a y  f s i l k f a .	s f r n a s g t k l q e q t l k l f y a e p v l s - s	k a m r l . I 9 y l a i y l s c p .	5800
AGAAGGCCTGGAGTGC  K a v e c s r r p w s a e g r g v l  c f a t s h l l g h i a s p r p t s  H B N G L L 1 1	AGGACATGACGCGGTTCTT  SCCCCA & k  PVLAPF LyLPQE  eqvqalf gtsagl srykrws  HHH H M HBXM I HA N B PIRB NAE F D ANGO P12 1 2 2121	deas ver ner ner ner ner ner ner ner ner ner n	MATTAGTTA:  . s i f n q . t i n i k i i	ITTAATT  i k c i k f i l n SHN FGS AIP	TTCGAATA  k l i s l . a y  f s i l k f a .	S f r n a s g t k i q e q t k i f y a e p v i s s BSB BHM SES SPN ACA PAL J1J 221	k a m r l . 19 y	5800
AGAAGGCCATGGAGTGC	AGGACATGACGCGGTTCTT  S C C C A K K  P V L B P F F  L y L F Q e  • Q V Q B L f  g C C C B K K  F V L B P F F  B F F F F F F F F F F  H HH H M HBXM  I HA N B PIRB  N AE F D ANCO  P 12 1 2 2121	deas ver reference deas ver reference deas ver reference deas ver reference deas ver reference deas ver reference deas ver reference de la ver ref	TTTATCATA  AAATTAGTTA:  f n q .  L i n i  . d i  n l .  k i i	ITTAATT  n . i k c i k f . f i l shi ffGS	TTCGAATA  k l i s l . a y  f s i l k f a .	S f r n a s g t k i q e q t k i f y a e p v i s s BSB BHM SES SPN ACA PAL J1J 221	k a m r l	
AGAAGGCCATGAGTGC	AGGACATGACGCGGTTCTT  S C C C A K K P V L A P F F L Y L F Q e  e Q V Q A L f g t S & g L S F Y K F W S  H HH H H HBXM I HA N B PINB H AE F O ANGO P 12 1 2 2121  GGGGGTTGGATTCTTGCCGG	deas ver representations of the second secon	TTTAATCAAT/	ITTAATT  n . i k c i k c i k . f . f i i l n SHN FGS AIP FACAAAT	TTCGAATA  k l i s l . a y  f s i l k f a .  TTHACGCCG	S f r n a s g t k i q e q i k k i f y a e p v i s s BSB BHM SES SPN ACA PAL J1J 221 /	k a m r l . I 9 y l a i y l s c p .  GGAATATGA	
AGAAGGCCATGAGTGC	AGGACATGACGCGGTTCTT  S C C C A K K  P V L B P F F  L y L F Q e  • Q V Q B L f  g C C C B K K  F V L B P F F  B F F F F F F F F F F  H HH H M HBXM  I HA N B PIRB  N AE F D ANCO  P 12 1 2 2121	deas ver representations of the second secon	TTTAATCAAT/	ITTAATT  n . i k c i k c i k . f . f i i l n SHN FGS AIP FACAAAT	TTCGAATA  k l i s l . a y  f s i l k f a .  TTHACGCCG	S f r n a s g t k i q e q i k k i f y a e p v i s s BSB BHM SES SPN ACA PAL J1J 221 /	k a m r l . I 9 y l a i y l s c p .  GGAATATGA	
AGAAGGCCATGGAGTGC	TCCTGTACTTGCGCCAAGAA AGGACATGAACGCGGTTCTT S C t C a k k P V l a P r r l y l r q e  e q V q a l f g t s a g l s r y k r m s  H HH H M HBXM 1 HA N B PINB N AE F D ANCO P 12 1 2 2121  GGCGCTTGGATTCTTGCCGG	deas ver to her some some some some some some some some	AMATTAGTTAT  S i  f n q .  L i n i  t . d i  n l . y  k i l	ittaatt  i k  c l k  f i l  f i l  f shh  FGS  AIP  HAZ  /  CACAAAT	TTCGAATA  k l i s l . a y  f s i t k f a .  TTTACGCG	S f r n a s g t k l q e q t l k l f y a e p v l s s BSB BHM SES SPN ACA PAL JIJ 221 / ACCTCGGTGTCCG	EGAATATGA	
AGAAGGCCTGGAGTGC  K a v e c : r r p w s a e g r g v l  c f a t s h l l g h l a s p r p t :  H B N G L 1 1  ACCGTAGAAGCCAATGAG  TGGCATCTTCGGTTACTC	AGGACATGAACGCGGTTCTT  S C C C a k k  P V i a P P P  I y i P q e  e q V q a i f g t s a g i S F y k F W S  H HH H M H HBXM I HA N B PINB N AE F O ANOC P 12 1 2 2121  GGCGCTTGGATTCTTGCCGG  CCGCGAACCTAAGAACGGCC	de a s v t r h r g r g i g  s s a d i l v l c r p r p m p  categoricatate  tragaagegttgatag i f a t m	AMATTAGTTAT  Sifn q. Lini Lini Lini Lini Lini Lini Lini Lin	i k k k k k k k f f k l n shn ssn ssn ssn ssn ssn ssn ssn ssn s	TTCGAATA  k l i s l . a y  f s i l k f a .  TTTACGCG	S f r n  a s g t  k l q e q  t l s e p v  l s s  BSB BHM SES SPN ACA PAL JIJ 221  ACCTCGGTGTCCG  TGGAGCCACAGGC  t s v s g	k a m r l . i g y l a i y l s c p .  GGAATATGA GGTTATACT	
AGCGTAGAAGCCAATGAC  H B H B H G L L 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AGGACATGAACGCGGTTCTT  S C C C a k k  P V l a P F F  l y l F q e  e q V q a l f  g t s a g l  s F y k F W s  H HH H M HBXM 1 HA N B PINB N AE F D ANGO P 12 1 2 2121  GGCGCTTGGATTCTTGCCGG  CCGCGAACCTAAGAACGGCC  F l d s C F  g a W i l a g	deas ver her of the second of	TTTAATCAATA  AAATTAGTTAT  f n q .  L i n i  L i n i  L i n i  CGCATCGGAGG  CCGTAGCCTCG  B S E i	ittaatt  i k  i k  c i k  f .  / i i  l m  SSHN  FGS AIP  MA2  /  TACAAAT  TACAAAT  T TACAAAT	TTCGAATA  k l i s l . a y  f s i l k f a .  TTTTACGCG  AAATGCGC	S f r n  a s g t  k l q e q  t k l f  y a e p v  l s - s  BSB BHM SES SPN ACA PAL J1J 221  / ACCTCGGTGTCCG  TTGGAGCCACAGGC  t s v s g  p r c p	k a m r l . I 9 y l a i r l s c p .  GAATATGA CCTTATACT	
AGCGTAGAAGCCAATGAC  H B H B H G L L 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AGGACATGAACGCGGTTCTT  S C C C a k k  P V i a P P P  I y i P q e  e q V q a i f g t s a g i S F y k F W S  H HH H M H HBXM I HA N B PINB N AE F O ANOC P 12 1 2 2121  GGCGCTTGGATTCTTGCCGG  CCGCGAACCTAAGAACGGCC	deas ver ner ner ner ner ner ner ner ner ner n	TTTAATCAATA  AAATTAGTTAT  f n q .  L i n i  n l .  k i l  GGCATCGGAGG  CCGTAGCCTCI  a s e i  h r s	ittaatt  i k i k c i k / i l i n  SHN FGS NIP NAZ / TACAAAT TACAAT TACAAAT TACAAAT TACAAAT TACAAAT TACAAAT TACAAAT TACAAT TACAAAT TACAAT  TTCGAATA  k l i s l . a y  f s i l k f a .  TTTACGCG	S f r n  a s g t  k l q e q  t k l f  y a e p v  l s - s  BSB BHM SES SPN ACA PAL J1J 221  / ACCTCGGTGTCCG  TTGGAGCCACAGGC  t s v s g  p r c p	k a m r l . I 9 y l a i r l s c p .  GAATATGA CCTTATACT		

FIGURE 3M

								Page 14
		_			_			
A L	H N	T H	SHH B			F K N N		
Ü	F	Å	UEE L			Üį		
1	3	î		1 1		H 1		
			//		·			
	TTGACGAAT						rcaaatggaggcbaagaac	
•							AGTTTACCTCCGCTTTCTTG	
a f	de, ltn	. sakn nprki	r'r p.	r l r g y g	tqyfr isise	k s = a :	kwrrkn inggert qmeakeh	q n
						-	k lh l r f f	
							e, fppslv	
							. i s a f s	
	_		***					
	s s		AM LA			M A	MH KF AN DO	
	P		UE UE			Ĥ	MF EX	
	1		11	ORF 3	TRANSLATION			OFLAVIN
				S	TOP	RBS-		ASE GENE
							TGGGAATCATATGAATATC	
AHIII	ITATAAACC	STITEGITTACT	GTTEGATECAGTA	AATGAAGTA	AAGATTAGTGTTTA	TAGTGTTTTTTCC	<b>TACCETTAGTATACTTATAG</b>	TATG
le i					•			
	níc	kok.t	s. v i	v f i	s n h 'k v	h k k a	<b>H e s v e v h</b>	•
							wesyeyh gnhmni	
. k	i fa	nqne	qarsf	t s f	litn	itkkd	gnhmni	iq
, k k k	ifa yio	nqne qtkæn	qars f klgh	tsf llh	litn f.sqi	i t k k d s q k r :	gnhmni ngii.is	i q y +
. k k k 	ifa ylo	nqne qtkmn lgfh	qars f kigh vi.tm	tsf llh .km	lith f.sqi +	itkkd sqkri 	gnhmni mgii.is hadysy	i q y + . v
, k k k 	ifa yio fiq ina	nqne qtkmn lgfh lgfks	qars f klgh vl.tm cald	tsf ilh .km nve	litn f.sqi +	itkkd sqkra 	gnhmni ngii.is hadysy spf.ifi	i q y + . v m
, k k k 	ifa yio fiq ina	nqne qtkmn lgfh lgfks	qars f klgh vl.tm cald	tsf ilh .km nve	litn f.sqi +	itkkd sqkra 	gnhmni mgii.is hadysy	i q y + . v m
, k k k 	ifa yio fiq ina	nqne qtkmn lgfh lgfks	qars f klgh vl.tm cald	tsf ilh .km nve	litn f.sqi +	itkkd sqkra 	gnhmni ngii.is hadysy spf.ifi	i q y + . v m
, k k k 	ifa yio fiq ina	nqne qtkmn lgfh lgfks	qars f klgh vl.tm cald	tsf ilh .km nve	litn f.sqi +	itkkd sqkri 	gnhmni ngii.is hadysy spf.ifi	i q y  . v m y
, k k k 	ifa yio fiq ina	nqne qtkmn lgfh lgfh lfwfs cvlif	qars fklgh vl.tm cald lsp.	tsfilh .km nve ks.	lith f.sqi tel.l nrivf k.dci H H B S B B	itkkd sqkri y,lfp ivff: dcfl	gnhmningii.is hadyay pf.ifi ipimhid	i q y  . v m y
, k k k 	ifa yio fiq ina	nqneqtkæn lgfh efwfs cvlif REH SCS APE	qars f klgh vl.tm cald lsp. H N	tsfilh .km nve ks.	lith f.sqi el.l nrivf k.dci  H H S S B B E O V	itkkd sqkra y.lfp ivff: dcfl  AF LN UU	gnhmningii.is hadyay ppf.ifi ipimhid EDESB H SDCFB I PETAV N	i q y  . v m y
, k k k 	ifa yio fiq ina	nqne qtkmn lgfh efwfs cvlif REH SCS	qars f klgh vl.tm cald lsp.	tsfilh .km nve ks.	lith f.sqi tel.l nrivf k.dci H H B S B B	itkkd sqkra y.lfp ivff: dcfl  AF LN UU	gnhmningii.is h s d y s y s p f . i f i i p i m h i d  EDE S B H SDC F B I PEI A V N 115 N 1 P	i q y  . v m y uff HN
kk kk if yf	ifa ylo fiq ina fyk	nqne, lgfh; lgfh; scvlifeREHSCSAPE	qars fklgh vl.tm cald lsp.  H N F	tsf (lh .km nve ks.	litn f.sqi rel.lin rivf k.dci H H B S B B E Q V	itkkd sqkre	gnhmningii.is hadysy pf.ifi ipimhid  EDESBH SDCFBI PETAV N 115 N 1 P	i q y . v m y HF HM AU 1H
k k if f y f l	ifa ylo	T K R N  L g f h  f w f s  C V L i f  R E H S C S A P E 1 1 1	qars fklgh vl.tm cald lsp.  H N F 1	tsfilhkm nve ks.	litn f.sqi rel.l nrivf k.dci HHB SBB EQV 1 2 1	itkkd sqkri y,lfp ivff: dcfl  AF LN UU 1H	g n h m n i m g i i . i s h s d y s y s p f . i f i i p i m h i d  EDE S B H SDC F B I PE1 A V N 115 N 1 P	i q y . v m y HF HN AU 1H
k k i'f y f l	ifa ylo	TERMINATION OF THE PROPERTY OF	qars fklgh vl.tm cald lsp.  H N F 1	t s filhkmn v e k s	litn f.sqi mel.l nrivf k.dci  M M B S B B E O V 1 2 1	itkkd sqkra y,lfp ivff; dcfl AF LN UU 1H	g n h m n i m g i i . i s h s d y s y s p f . i f i i p i m h i d  EDE S B H SDC F B I PET A V N 115 N 1 P	i q y y  HF HN AU 1H
AAGGAUTICET	ifa ylc fiq ina fyk	T CCATGTCCAGAAT	q a r s f k l g h v l . t m c a l d l s p .  H N F 1  MAATCGGAATCGT	t s f	l i t n f . s q i m r i v f k . d c i  H H B S B B E Q V 1 2 1  ATTTAATGATTTTA	itkkd sqkra y.lfp ivff: dcfl  AF LN UU 1H TTACGAGCAAGCT	g n h m n i m g i i . i s h s d y s y s p f . i f i i p i m h i d  EDE S B H SDC F B I PET A V N 115 N 1 P / GCTGAGCGGAGCAGAAGATG	i q y . v m y HF HN AU 1H CGCT
AAGGAUTTCCTT	ifa ylc fiq ina fyk	T CCATGTCCAGAAT	q a r s f k l g h v l . t m c a l d l s p .  H N F 1  MAATCGGAATCGT	t s f i l h . k m n v e k s .  AGTAGGAAG	l i t n f . s q i m l . l n r i v f k . d c i  H H B S B B E Q V 1 2 1  ATTTAATGATTTTA TAAATTACTAAAAT.	itkkd sqkra y.lfp ivff: dcfl  AF LN UU 1H  TTACGAGCAAGCTT	g n h m n i m g i i . i s h s d y s y s p f . i f i i p i m h i d  EDE S B H SDC F B I PET A V N 115 N 1 P / GCTGAGCGGAGCAGAAGATG	i q y . v m y HF HN AU 1H CGCT+ 4200
AAGGAUTTCCT	ifa ylc fiq ina fyk	TEGTACAGGTCTAGAAT	q a r s f k l g h  v l . t m c a l d l s p .  H N F 1  MAATCGGAATCGT  TITAGCCTTAGCA	t s f l h . k m n v e k s .  AGTAGGAAG TCATCCTTC	litn f.sqi m.rivf k.dci  H.H.B.S.B.B.E.O.V.1 2.1  ATTTAATGATTTTA TAAATTACTAAAAT.	itkkd sqkri y.lfp ivff: dcfl  AF LN UU 1H  TTACGAGCAAGCTI  TTACGTCGTTCGAI	gnhmni mgii.is hadysy pf.ifi ipimhid  EDE SB H SDC FB I PE! AV N 115 N 1 P / CCTGAGCGGAGCAGAAGATG	i q y y  HF HN AU 1H  CGCT 4200
AAGGA	ifa y to fiq ina fy k  MATTTAGTT  TTAMATCAM	T CCATGTCCAGAAT	q a r s f k l g h v l . t m c a l d l s p . H N F 1 1 MAATCGGAATCGT	t s f i l hk m n v e k s  AGTAGGAAG TCATCCTTC s r k v 9 re d	I i t n f . s q i t e l . i n r i v f k . d c i  H H B S B B E O V 1 2 1  ATTTAATGATTTTA  TAAATTACTAAAAT. i . f y f n d f i l m i l	itkkd sqkre y,lfp ivff: dcfl  AF LN  UU  TTACGAGCAAGCTE  TTACGAGCAAGCTE  T skl lrasc	gnhmni mgii.is hadysy pf.ifi ipimhid  EDE SB H SDC FB I PEI AV N ITS N 1 P CCTGAGCGGAGCAGAAGATG  CGACTCGCCTCGTCTTCTAC Bersrc L Sgaeda . aeqkm	i q y . v m y HF HN AU 1H CGCT
AAGGAATTICCT	fiq ina fyk	T CCCATGTCCAGAAT	q a r s f k l g h v l . t m c a l d l s p .  H N F 1  MAATCGGAATCGT  TITAGCCTTAGCA  T	t s f i l h . k m n v e k s .  AGTAGGAAG TCATCCTIC	litn f.sqi mel.l nrivf k.dci  H H B S B B E O V 1 2 1 ATTTAATGATTTTA  TAAATTACTAAAAT.	itkkd sqkri y,lfp ivff; dcfl  AF LN UU 1H TTACGAGCAAGCTI ,	gnhmni mgii.is hadysy pf.ifi ipimhid  EDE SB H SDC FB I PEI AV N 115 N 1 P CCTGAGCGGAGCAGAGATG  CGACTCGCCTCGTCTTCTAC a e r s r c i s g a e d a . a e q k m	i q y y HF HN AU 1H CGCT
AAAGGAU	fiq ina fyk	T CCATGCCCAGAAT	q a r s f k l g h v l . t m c a l d l s p .  H N F 1  MAATCGGAATCGT  FITTAGCCTTAGCA  D F F F C S & S .	t s f t l h . k m n v e k s .  AGTAGGAAG TCATCCTIC s r k v 9 r . e d	litn f.sqi mel.l nrivf k.dci  HHBSBBEOV 1 2 1 ATTTAATGATTTTA  TAAATTACTAAAAT. i.fy fndfi ldmil	itkkd sqkra y,lfp ivff; dcfl  AF LN UU 1H TTACGAGCAAGCTT  ATGCTCGTTCGAI yeqa; tskl lrascca	gnhmni mgii.is hadysy pf.ifi ipimhid  EDE SB H SDC FB I PEI AV N ITS N 1 P CCTGAGCGGAGCAGAAGATG  CGACTCGCCTCGTCTTCTAC Bersrc L Sgaeda . aeqkm	i q y  HF HN AU 1H EGGT 4200 GGGA

FIGURE 3N

						Page 1
MD N A		HE AS H H	8	F	s s	
8D L C		LC PC 1 H		N	Q F	
OE A C		AR YR N A		U	O A	
21 3 1		42 11 P 1	1	н	1 N	
GCTCAGACATGGCGTAGAC		CTTGGGTTCCAGGCGC			MATGGCGGAMACMAMA	/300
					TTTACCGCCTTTGTTTTTT	4300
lrhg v d s d m a . t	tndid va qatla w	wypga lgfqah	feip lkyr	faakk llrk:	k w r k q k n	
a.vhrl	clhcqhp	k p e l r	mqfv	tqqsf		
Slepts qesmay	v f s m s t civníh	eqtgpa spnwad	nsig	naaf: ksrf	f <b>ias</b> v f f f h r f c f f	
	NM R	H E	8	AF	THH S F	
		H C	В	LN	HIN F N	
		A 0	V	w	ANA A U	
	2 1 P	1 D	1	1H	1P1 N H /	
cyyhi ydaiiti mmilsh	ghchqragtagtattatt ghchqri gtvirg walssea	endtlr ndtlr atthyd qrhit	TAATACAGACGT lclq yvcn imsa	SCKI SCKI BAK K	giaqaan asrkqq	
ihhm		r l s v n r	nhre s.tq	hiqi lsaaf	icralli pmacaa	
B R	H	т	T	CHN		
8 5	A	A	A	FPA		
V A 1 1	E •	9	9	RAE		
1 1	3	1	1	221		
ACACTACTGGTGTACCTGT	ATETTTGGAATTGTAACA	CTGAAAACATCGAACA	GCTATCGAGCG	TGCCGGCACAAAA	CGGGCAACAAAGGTGTAGA	
,	·, · <del>-</del> ,	*****.		+, + -	GCCCGTTGTTTCCACATCT	4500
	TABOACCI IAACAI I GI ;	BACTITIO INDELIGI	LUA I AGC I CGC	ACGGCCG1G1111C	GCCCGITGITTCCACATCI	
hywctci ttgvpv tlivyis	ilwacan ifgivtt stel.q	enieq	a i e r	agtka	gnkgvd	
	·····	+,+,			,	
fvvptgt	.rqfqll mkpitv idksnye	v s f m s c	aisr	a p v f	aplipts	

FIGURE 30

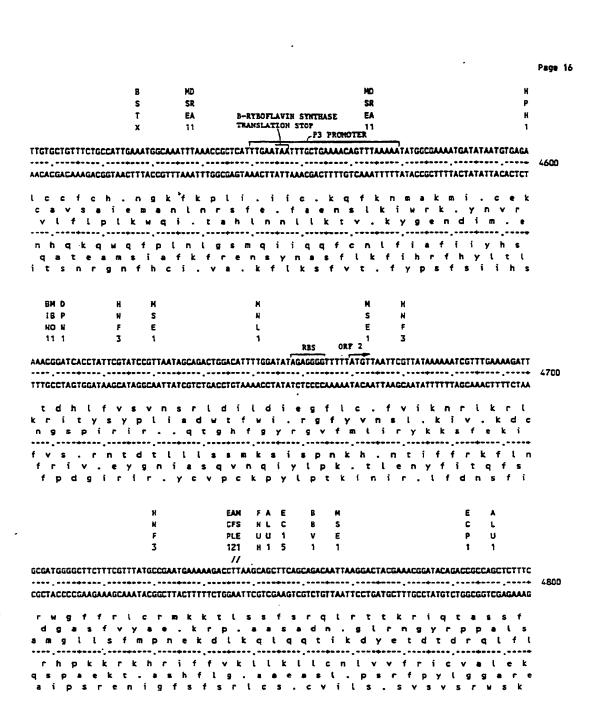


FIGURE 3P

M N E H T P M H D XMD HB H N N C N A L N N D HBP PI N L L R F Q E L F E COM AN F 1 1 V 1 1 1 1 1 1 211 21 1  TTTGGAAAGAGGACGAGGATACGTCGGAGCAATCGGAGTCGAAAAAAAGGATTCTGAGGTTGAGATCCGGCATATCAGTGTGAAATCA  AAACCTTTCTCCTGCTCCTATAGCAGCCTCGTTAGCCTCAGCTTTTTTTCCTAAGACTCCAACTCTAGGCCGGATATCAGTGGAAAC  f g k r t r i s s e q s e s k k r i l r l r s g i s v . i l e r g r g y r r s n r s r k k g f . g . d p a y q c e s w k e d e d i v g a i g v e k k d s e v e i r h i s v n p  k p f l v l i d d s c d s d f f l i r l n l d p m d t h i k s l p r p y r r l l r l r f f p n q p q s g a y . h s d q f s s s s i t p a i p t s f f s e s t s i r c i l t f g  H S ND F SR D CS L S L F A EA K AA A E U 1 N 11 1 11 1 11 4 1 1  CAAGGAATCGGAAAACAGATGATGGATGCTTTAAAGCATTTATTCAAAACGCAAGGTTCCAAATGAATTAACGCAGGGTTT  CAAGGAATCGGAAAACAGATGATGGATGCTTTAAAGCATTTATTCAAAACGCAAGGTTCCCAAATGAATTAACGCAGGGTTT  CAAGGAATCGGAAAACAGATGATGGATGCTTTAAAGCATTTATTCAAAACGCAAGGTTCCCAAATGAATTAACGCAGGGTTT  CAAGGAATCGGGAAAACAGATGATGGATGCTTTAAAGCATTTATTCAAAACGCAAGGTACTGGTTCCAAATGAATTAACGCAGGGTTT	GAAGAGTAGCGGT  L i a i f s s p s h r h
H C NAL NND HBPPI N L R F G E L F E COM AN F 1 V 1 1 1 1 1 1 1 211 21 1  TTGGAAAGAGGACGAGGATATCGTCGGAGCAATCGGAGTCGAAAAAAAGGATTCTGAGGTTGAAGTCCGGCATATCAGTGGAATCA  AACCTTTCTCCTGCTCCTATAGCAGCCTCGTTAGCCTCAGCTTTTTTTCCTAAGACTCCAACTCTAGGCCGTATAGTCACACTTAGT  g k r t r i s s e q s e s k k r i l r l r s g i s v . i l e r g r g y r r s n r s r k k g f . g . d p a y q c e s w k e d e d i v g a i g v e k k d s e v e i r h i s v n p  k p f l v l i d d s c d s d f f l i r l n l d p m d t h i k s l p r p y r r l l r l r f f p n q p q s g a y . h s d q f s s s s i t p a i p t s f f s e s t s i r c i l t f g  H S ND F SR N M A N F SR O CS L S L 1 N 11 1 1 1 1 1 1 1 1	GAAGAGTAGCGGT  L i a i f s s p s h r h  r r m a k e d g j e . r w
F G E L F E COM AN F V 1 1 1 1 1 1 1 1 211 21 1  TTGGAMAGAGGACGAGGATATCGTCGGAGCAATCGGAGTCGAAAAAAAGGATTCTGAGGTTGAGATCGGCCGTAATCAGTGTGAATCA  ACCTTTCTCCTGCTCCTATAGCAGCCTCGTTAGCCTCAGCTTTTTTTCCTAAGACTCCAACTCTAGGCCGTATAGTCACACTTAGG  g k r t r i s s e q s e s k k r i l r l r s g i s v . i l e r g r g y r r s n r s r k k g f . g . d p a y q c e s w k e d e d i v g a i g v e k k d s e v e i r h i s v n p  P f l v l i d d s c d s d f f l i r l n l d p m d t h i k s l p r p y r r l l r l r f f p n q p q s g a y . h s d q f s s s s i t p a i p t s f f s e s t s i r c i l t f g  H S NO F SR N H A N F SR O CS L S L I N I N I N I N A I N I N I N I N A I N I N I N A E U I N II N II N II N II N II	GAAGAGTAGCGGT  L i a i f s s p s h r h  r r m a k e d g j e . r w
TEGAMAGAGGACGAGGATATCGTCGGAGCAATCGGAGTCGAAAAAAAGGATTCTGAGGTTGAGATCCGGCATATCAGTGTGAATCAGCTTTTTTTCCTAAGACTCCAACTCTAGGCCGATATCAGTGTGAATCAGACTCTCTGCTCCTATAGCAGCCTCGTTAGCCTCAGCTTTTTTTCCTAAGACTCCAACTCTAGGCCGATAAGTCAACACTTAGG  g k r t r i s s e q s e s k k r i l r l r s g i s v . i l e r g r g y r r s n r s r k k g f . g . d p a y q c e s w k e d e d i v g a i g v e k k d s e v e i r h i s v n p  p f l v l i d d s c d s d f f l i r l n l d p m d t h i k s l p r p y r r l l r l r f f p n q p q s g a y . h s d q f s s s s i t p a i p t s f f s e s t s i r c i l t f g  H S MD F SR N M A A B M F SR D CS L S L I N II I I I I I I I  1 1 1 1 1 1 1 1 1 1 1	GAAGAGTAGCGGT  L i a i f s s p s h r h  r r m a k e d g j e . r w
TGGAAAGAGGACGAGGATATCGTCGGAGCAATCGGAGTCGAAAAAAAGGATTCTGAGGTTGAGATCCGCCATATCAGTGTGAATCC  ACCTITCTCCTGCTCCTATAGCAGCCTCGTTAGCCTCAGCTTTTTTTCCTAAGACTCCAACTCTAGGCCGTATAGTCACACTTAGG  g k r t r i s s e q s e s k k r i l r l r s g i s v . i e r g r g y r r s n r s r k k g f . g . d p a y q c e s w k e d e d i v g a i g v e k k d s e v e i r h i s v n p  p f l v l i d d s c d s d f f l i r l n l d p m d t h i k s l p r p y r r l l r l r f f p n q p q s g a y . h s d q f s s s s i t p a i p t s f f s e s t s i r c i l t f g  H S ND F SR N M A N F SR O CS L S L I N I I I I I I I I I I  1 1 1 1 1 1 1 1 1	GAAGAGTAGCGGT  L i a i f s s p s h r h  r r m a k e d g j e . r w
ACCTITETECTICCTCCTATAGEAGCCTCGTTAGCCCTCAGCCTTAGCCCGTATAGTCACACTTAGGCCGTATAGTCACACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACTTAGGCCGTATAGTCACACACTAGACACACAC	GAAGAGTAGCGGT  L i a i f s s p s h r h  r r m a k e d g j e . r w
gkrtrisse asekkrilrtrsgisv.i ergrgyrrsnrsrkkgf.g.dpayaces wkededivgaigvekkdseveirhisvnp  pflvliddscdsdfflirlnldpmdthi kslprpyrrltrlrffpnapasgay.hsd afssssitpaiptsffsestsirciltfg  HS NOF SRN MAA BN FSRO CSL SL FA EAK AAA E U 11 N 11 1 11 1	ttiai fssp shrh rrma kedg Je.rw
wkededivgaigvekkdseveirhisvnp  Pflvliddscdsdfflirlnldpadthi kslprpyrrllrlrffpnqpqsgay.hsd qfssssitpaiptsffsestsirciltfg  HSNOFSROCSLSL FAEAKAAAEU  1 N 11 1 11 4 1 1	fssp shrh rrma kedg je.rw
Pflvliddscdsdfflirtnldpmdthikslprpyrrltrlrffpnqpqsgay.hsdqfssssitpaiptsffsestsirciltfg  HS MOF SRN MAA NF SRO CSL SL FA EAK AAA E U 1 N 11 1 11 4 1 1	shrh rrma kedg je.rw
Pflyliddsedsdflirlnldpmdthikslprpyrrltrlrffpnqpqsgay.hsdqfssssitpaiptsffsestsirciltfg  HS NOF SRN MAA NF SRO CSL SL FA EAK AAA E U 1 N 11 1 11 4 1 1	rrma kedg e.rw
ks lprpyrriirffpn qpq sgay.hsd qfs sssitpaipts ffs est sircilt fg 	rrma kedg je.ru
qfsssitpaiptsffsestsirciltfg  HS MDF SRN MA  NF SR O CS L S L  FA EAK AAA E U  1 N 11 1 11 4 1 1	kedg Je.ru AT M
H S MD F SR N M A N F SR O CS L S L F A EA K AA A E U 1 N 11 1 11 4 1 1	3 Ф . г н АТ Н
N F SR 0 CS L S L F AA A E U 11 11 4 1 1	
N F SR 0 CS L S L F AA A E U 11 11 4 1 1	
N F SR D CS L S L F A EA K AA A E U 1 1 1 1 1 1 1 1 1	
F A EA K AA A E U 1 N 11 1 11 4 1 1	
1 N 11 1 11 4 1 1	UQ E
	21 2
ACCRETCOLARADAMANA	, -
esenr.wml.siyskrkywfqmn.rraf rnrktddgcfkafiqnastgsk.inaelf gigkqmmdalkhlfktqvlvpneltqsf	rtls
sdsflhhie im a fall a million in a million in a million i	*****
pipfciis akfcknlvet stgfsnvelk	ksrq
E SM D MM A BMD AR F	
C 10 B	MB
1 NON LE U LON UNO	AS
5 11 1	EA
STOP / TRANSCRIPTION TERMINATOR	21
GTCAGCAGGATCAAGACATTTCATACAATAATTAAGCAGGCTGTGATCACCCCCCCC	—
CAGTCGTCCTAGTTCTGTAAAGTATGTTATTAATTCGTCTCCGACACTAGTCAGACGACGAAAAAAAA	AGTGCAAGTGC
OPP 1 TRIMET LETON COOR	
v s rikt fhtiik q r las v s a f f s a f y f f f	
vsriktfhtiikqrl.svsaffsafyff sagsrhfiq.lsrgcdqsllfflrsisf:	Srsr
199 do disynn, a eaviele 445 m	S r S r h v h a
qqdqdisynn.aeavislcfflrsisf	srsr hvhg
qqdqfsynn, aeavislefflrsisf	s r s r h v h g
qqdqdisynn.aeavislcfflrsisf	s r s r h v h g +

FIGURE 3Q

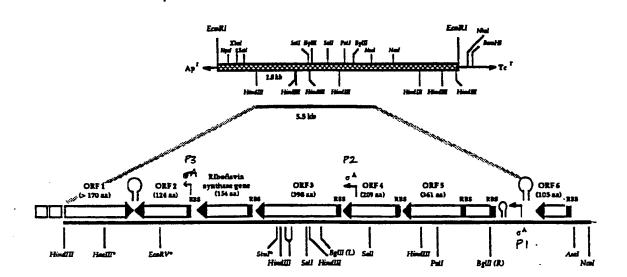
			_							Page 18
									•	*444
A MA	FM D B		T	CB			F	SHA		TMH ABN
HA A	OS P I	P	A	DB		GS	N	NLC		ODF
	KO N N		٥	EV		IP.	Ü	AAC		113
22 2	11 1 1	1	1	11	ī	AZ	н	131	•	
/				/		/			TTTTTCCTCCA1	TER
GATGACGTCAGT	CCGATCCC	CAAACGGTGTT	TGTCGATAAGA	ATATGTTGCTG		AL 161		CCCATGTATACTTT		+ 5200
				TATACAACCAC	TOAC	TEAC	CCACGG	GGGTACATATGAAA	AAAAAGGACGT/	MGC
ddvs mtsv	pip rsr	qtv f kral	vdkk siri	yva e nall icci.	: c 5 &	t g	i p	pcilf phvyf pmytf	fpat fflh fsci	f d s
sstl	a i a	cvtn	it sl	fy ta	s h	V	P \$ {	, , , , ,	K K 9 B	n
ivdt	trd:	rlch	kdil	fins	L		Pq	gutyk	KKFC	•
phr.	d 2 0	a f p t	d t A e	i h q	ĮŢ	c q		gmyvk	. = 4 :	
D B SNI	N 8	×			В		ST			T
P I PSI		N N			\$		PA			A
N N HP/		Ü			M		00			0
1 1 183		1			2		11			1
1	,									
ATCCTGCATGC	TTECTCCAG	TTTCTCATCTT1	rgattggcagta	TAATGCTTTTA	TAGGC	AGAGA	CGGTTT	COATTTGTTCGTAM	CCGATTGCATA	AGTT
						<b>+</b>		,		3300
TAGGACGTACG	AAGGAGGTC	MAGAGTAGAN	ACTAACCGTCAT	ATTACGAAAAT.	ATCCG	TCTCT	CCCANA	GCTAAACAAGCATT1	GGCTAACGTAT	TCAA
pac	fiq	flif	q m d A	nafi	g	r d	gf	dlfvn		, T
ilha	<b>S S S</b>	fasl	្ត្រែ s	n i i		• 1	; V B	ics.1		3 3
scal	P P V	shi.	. L . v	. c † y	r 0	F	г т 	rfvrk	p ,	
					 i n	1 .	ı p k	s k n t	rnc	l n
s g a n	K F W	13 F M &	8 Q E ) 		٠, ٠,		, p -	e i q e y	v s q m	ι
1 7 6 8				v h k .	΄, '	e l	ר ח	rntri	giaj	/ t
0 Q m			4	,		•			-	
GCX	BNMD	H D	TB NS	Ī	S H	BH	M	F	BSBNSH	
DFA		B D	AB CCI	ì	F N	SP	8	0	SESCCP	
IRE	LAON	0 E	QV 1R/		A F		0	ĸ	ACAIRA	
213	1311	2 1	12 117	2	N 1	22	Z	1	J1J112	
/	11		//	7					/ //	
CGAGCAAACGG	CCATGATC	MAGCCCTAAGTC	TTCGACTGCCC	GTGTTCTGCTT	GMC	MTCC	GCATGCT	GTTCGCCATCAGTC	TTTTTGCCCCG	561G1 5600
			,				-,			LUTCY
GCTCGTTTGCC	EGGTACTAG	TTCGGGATTCAG	AAGET GACGGG	CACAAGACGA		IIAOO	CC I ACC	LEAAGEGGTAGTEAG		<b></b> .
				v 4 e l			d a	v r h q s	fcpg	c
e q t 1				C S A	'	i F	m l	faisl	fap	<b>.</b> v
3 K r	p. s		rle	a v l l			g C (	s p s v	flpr	l y
						.+	-,			
s c v	a m i	lg t	k s 0 a	tnql	k f	f g		t r w . d	kqg	Pq
ellr	g h d	lgic		r h e a	q l	i	r i :	s nam l	rkag	a t
rafi	9 W S	.arl	rrsg	ptrs		s d	рh	qegdt	k k g r	S
			-							

FIGURE 3R

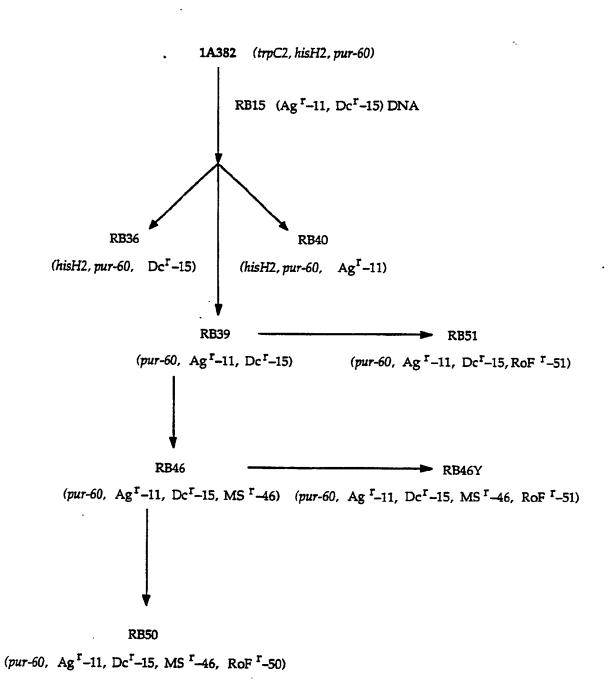
																																			Page
											A													н	8	XM	•		M		s				
											L													8	1	HB	•		В		5				
											U													0	N	00	ŧ		0	i	P				
											1													2	1	21	1		2		1				
TCTGC	·rti	rto	TC	l TC				~, ~	***		^~															/						. <b>.</b>			
TCTGC									166			-	AI,		GC:C	CAU		LATA	GCC	IIT.	TTC	GT	TTC	1110	3 <b>G</b> C	56	LTC	TTC	CTT	CCA	AT/	LTTC	TT(	CT	5500
AGACO	GLA	GAC	ACI	AC	TAT	LTT	rca	GTG	LCG	TT	CCA	CT	TAT	TT	cee	CTC	cct	TAT	CC	44		CA	146							+ 	·		•	-+	5500
						•						•••	•	• • •						~~	~~	~~	~~	~~:			~~	~~	-	401		AAG	•	UA.	
i p	s	٧		3	ı	. :	3 (	h (	3	k	ί	n	k	: ;	•	h	P	i	2	f	\$	f	ι	w	٦	i	i	f		<b>D</b>	i	f	f	_	
c	ŧ	ŧ	•	•	Y	k	2	τ		2			ī	k	P	t	q		г	f	r	. 1	F	fs	3	9	5	5	f	a	Y	2	2		
	f	c	đ	đ	i	k	P	ı	q		2	e	•	\$	P	P	n	. s	٧	, ,	f	v	8	ŧ	2	ď	ι	P	8	ก	i	l		ι	
		•••	٠.٠																																
_	_	_	_ `		•					•:		-+	•••	•		-+-					٠	•			• • •			,	•		•••	-,-	•••	-+	
r	g	e	t	i	i	Y	ι	¥	q	l	5		f	t		w	8	i	š	n	•		1		1	г	i	k	۲	œ	j	n	k	-+	
u d	9	e r	t h	i h	i y	γ l	l a	۷ ۲	۹ •	ı	<b>.</b>	q	f	l f		u v	8	i y	e F	n	•	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
ı q	9	e r	t h	i h	i y	γ l	l a	۷ ۲	۹ •	ı	<b>.</b>	q	f	l f		u v	8	i y	e F	n	•	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
ı d	9	e r	t h	i h	i y	γ l	l a	٧ ٧	۹ •	ı	<b>.</b>	q	f	l f		u v	8	i y	e F	n	•	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
, d e e	g k	e r	t h	i h	i y	γ l	l a	٧ ٧	۹ •	ı	\$    a	q s	f	l f		u v	8	i y	e F	n t	k	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
r 1 q e a	g k	• •	t h	i h	i y	γ l	l a	٧ ٧	۹ •	ı	<b>.</b>	q s	f	l f		u v	8	i y	e F	n	k k	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
, d e s	g k	e	t h	i h	i y	γ l	l a	٧ ٧	۹ •	ı	S a N	q s	f	l f		u v	8	i y	e F	n t	k k	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
r q e a	g k	e C NB LS	t h	i h	i y	γ l	l a	٧ ٧	۹ •	ı	5 8 N L	<b>q</b>	f	l f		u v	8	i y	e F	n t H	k ALU	r	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
r q e a	g k	e P NB LS AM 31	t h	i h : 1	i y	Υ ι	l •	¥ ; 1	q •	c	S N L A	q s	f i y	t f	<b>9</b>	<b>u</b> 9	9 9	i Y		1 H I H 3	k ALU1	r t	i k	r c	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
e a	g k k	e r q	t h   S	i h	y y	Y L	l a F	e cre	q a	C GCJ	S A N L A 3	9 8	f i y	t f	g L	V V	g g	i Y	ATA	TAA	k A L U 1	: r r t	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
e a	g k	e P	t h i s	i h	y i i	Υ   		W V 3 1	q	GCJ.	A A A A A A A A A A A A A A A A A A A	9 5	f i y	[   f 	g l	V V	g g	i y	ATA	TAA	k ALU1	TT	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
n q e a	g k	e P P P P P P P P P P P P P P P P P P P	t h i s	i h	y i i	Υ   		W V 3 1	q	GCJ.	A A A A A A A A A A A A A A A A A A A	9 5	f i y	[   f 	g l	V V	g g	i y	ATA	TAA	k ALU1	TT	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
e a	g F k	MB LS AM 31	t h s	i h	y y TTA	AAT	CCC GGG	GTG	q a i i i i i	GC.	H L A 3	Q S	f i y	CAA	9	M V V V V V V V V V V V V V V V V V V V	g g	i Y L	ATA	TAA	k A L U 1	TT	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
e a	g r k	MB LS AM 31	t h s TAA	i h ::::	i y i i	AAI	CCC GGG	V V S S S S S S S S S S S S S S S S S S	q a i i i i i	GC.	N L A 3	q s	f i y	CAA	9	W V P	g g	i y	ATA	TAAA	k A L U 1	TT AA f	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
ATTT	g F k	er q	t h s	i h:::	i y i i	Y L TTA	CCC GGG	V V S S S S S S S S S S S S S S S S S S	q a	GCJ CG1	N L A 3	q s	f i y	CAA	i g	W V	g g	i y l	ATA L TAT,	TAA	k A L U 1 GC	TT AA f	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	
e a	GTAI	MB LS AM 31 CCC	t h s	i h::::	AAT	AAI TTA	GGG P	SCAC	q a	GCJ CG1	H L A 3	q s	f i y	CAME	ing of the second	d AGA	g g	i y	ATA TAT.	n i t H I N 3	c k A L U 1 GC . CG	TT AA f	k e	k k	Į P	r 0	i d	k e	r k	g	i Y	n ' e		•	

FIGURE 3S

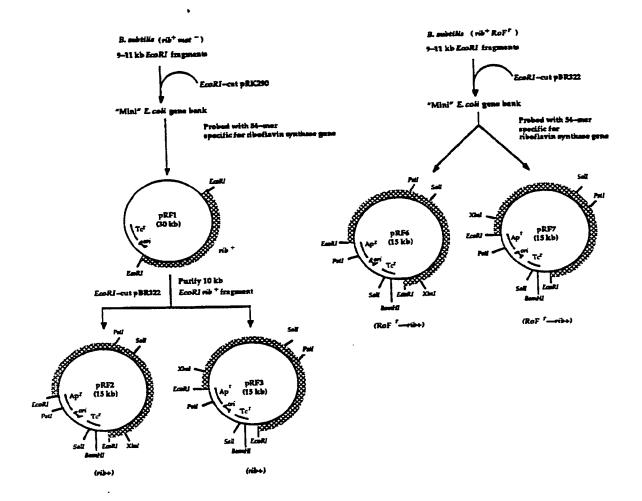
# FIGURE 4.



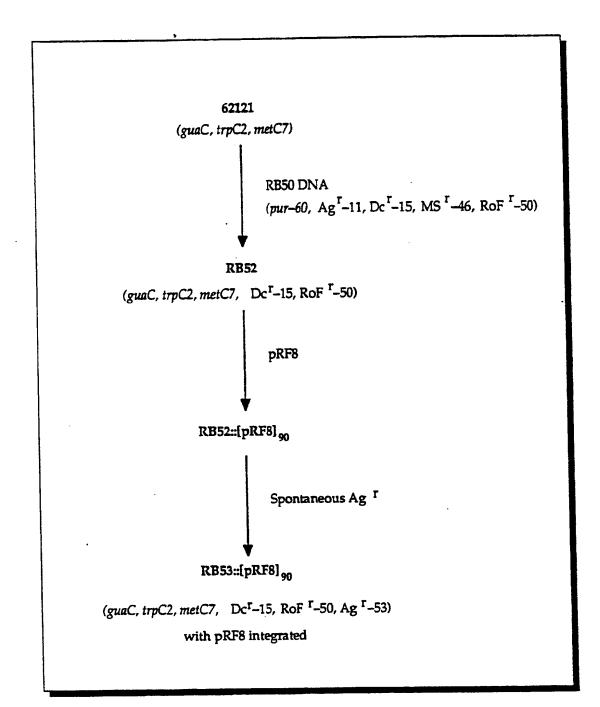
### FIGURE 5.



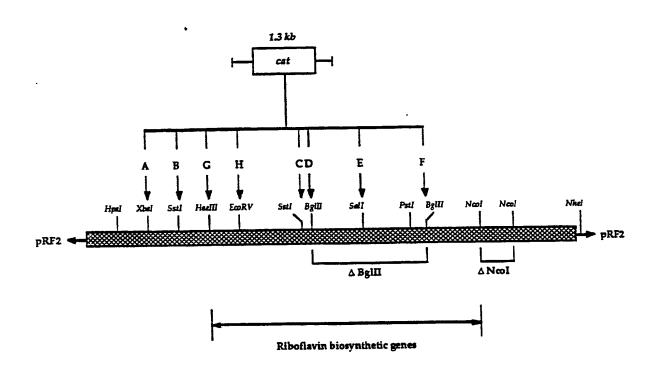
## FIGURE 6.



### FIGURE 7.



# FIGURE 8.



### FIGURE 9.

Within 5'

Upstream from

 $\Delta G = -20 \ kcal/mol$ 

At 3' end of

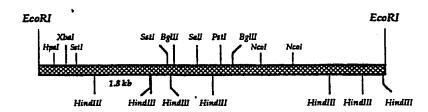
 $\Delta G = -16.5 \ kcal/mol$ 

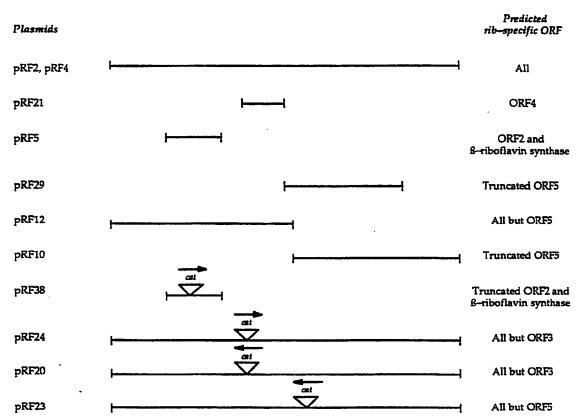
ribP1	leader mRNA	rib operon
A CGAAAGGGTGATGTT	T A T T A A T T A A T A A T C C C C C C C C C C C C C C C C C C C	G CAGTCTCTGG.CG.CG.CG.CG.CG.CG.CG.CG.CG.CG.CG.CG.C
A·T ~ AA·TG(T <sub>5</sub> )	G·C ~AAA·T(T <sub>5</sub> ) ~	A•T ~ TA•T(T <sub>6</sub> ) ~
#708 #742	#1034   #1062	#5037 #5064

59

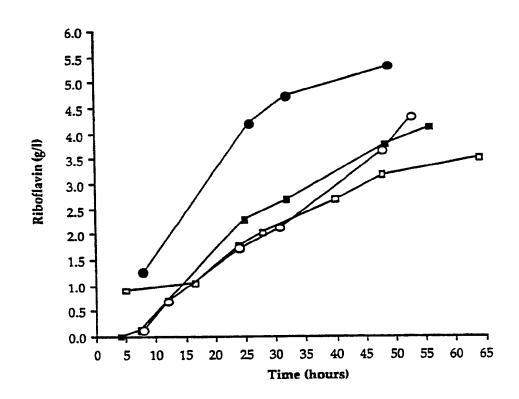
 $\Delta G = -26 \ kcal/mol$ 

## FIGURE 10.

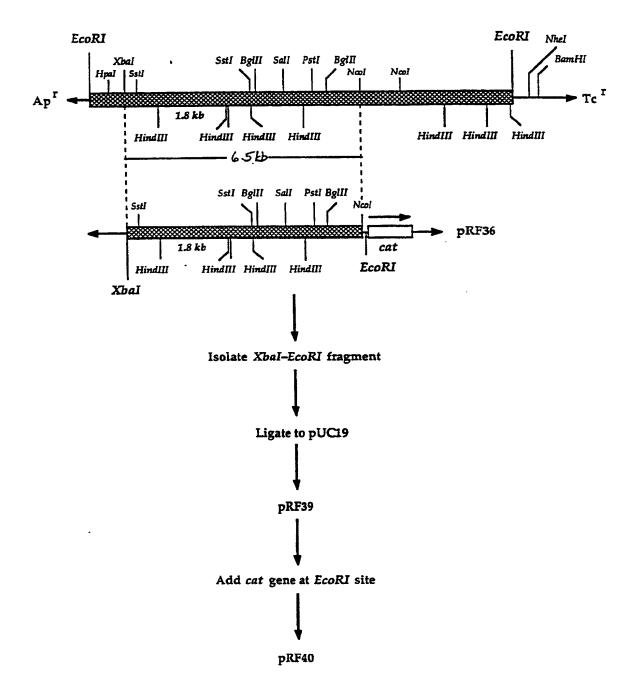




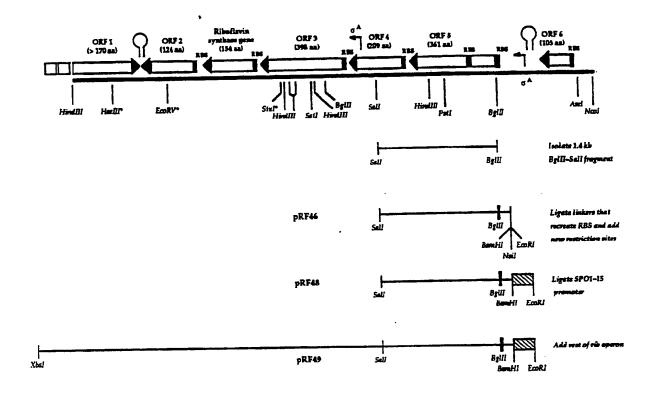
# FIGURE 11.



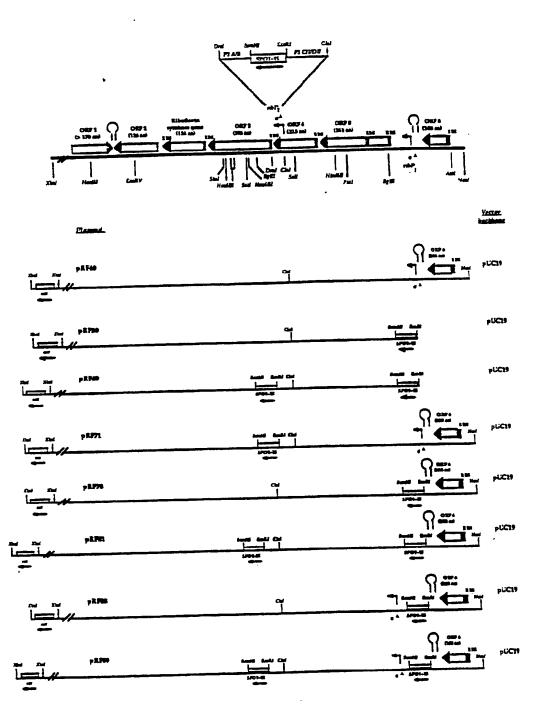
### FIGURE 12.



## FIGURE 13.



# FIGURE 14.



## FIGURE 15.

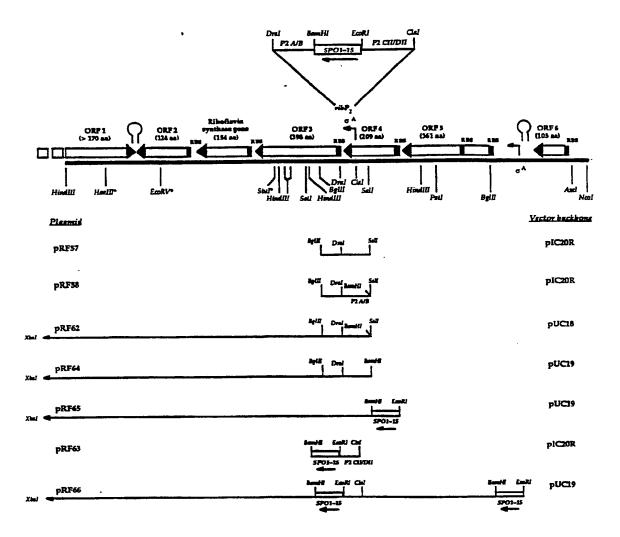
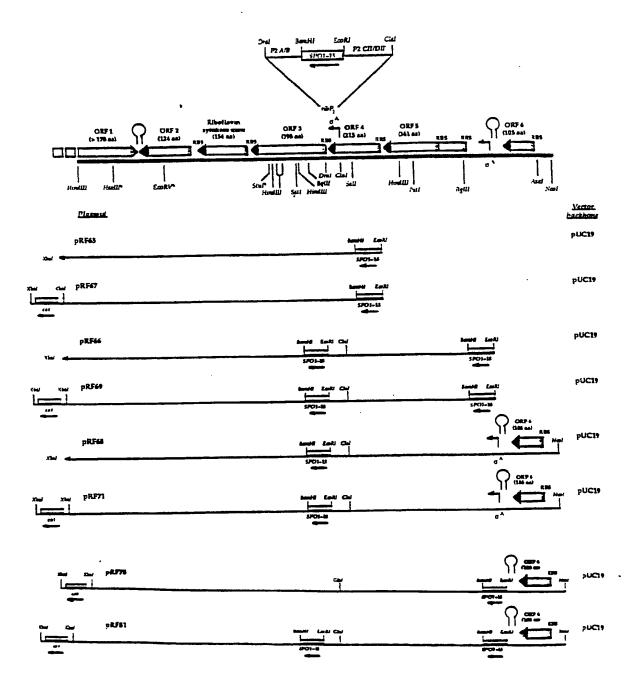


FIG. 16



F16.17

GAATTCCCGGGATCC

55-mir

TGATTGCAAGCCTTTCGGATCGAAGCGTGATGTTTTGTTTTCTCATTGTAAAAGCATCGCTATAATT

•

## FIGURE 18

RB-5	AATTCATGCATGGATCCGACGGTAAATAAC AAAAGAGGGGAGGG
RB-6	GATCTAAGGCCAGCTTCATATAATACTCTT CCATTTGTTTCCCTCCCCTC
P2-A	TCGACGGATCCTTTTAGAGAGGAAGATTTG CATGTTTCATCCGATAGAAGAAGCACTGGA CGCTTT
P2-B	AAAGCGTCCAGTGCTTCTTCTATCGGATGA AACATGCAAATCTTCCTCTCTAAAAGGATC CG
P2-CII	CGATTTTTGCATAAAGCCAATGAAAATAAG ACCCAACAAACCATTACAAAAGCCTTCTTA AGCGAAAACGGCTTTTAG
P2-DII	AATTCTAAAAGCCGTTTTCGCTTAAGAAGG CTTTTGTAATGGTTTGTTGGGTCTTATTTT CATTGGCTTTATGCAAAAAT



## **EUROPEAN SEARCH REPORT**

Application Number

EP 90 11 1916

Category		indication, where appropriate,	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
D,Y	FR-A-2 546 907 ( NAUCHNO-ISSLEDOVATE GENETIKI I SELEKTSI MIKROORGANIZMOV (VN * Whole document *	/SESOJUZNY ELSKY INSTITUT II PROMYSHLENNYKH	1-5,8- 13,15	C 12 N 15/52 C 12 P 25/00 C 12 N 1/20 / (C 12 N 1/20 C 12 R 1:125
Y	subtilis riboflavio	293, abstract no. Ohio, US; P.M. "Expression in ells of the Bacillus operon in hybrid PR2", & DOKL. AKAD.	1-5,8-13,15	C 12 R 1:19)
Y	no. 86284h, Columbu JOMANTIS et al.: "N	?, page 214, abstract is, Ohio, US; J. Method for the	1-5,8- 13,15	
	successive cloning Bacillus chromosome NAUK SSSR 1982, 264 * Whole abstract *	es", & DOKL. AKAD.		TECHNICAL FIELDS SEARCHED (Int. CL5)
Y	CHEMICAL ABSTRACTS, 11th April 1983, pa 120613x, Columbus, PANINA et al.: "Clo riboflavin biosynth Bacillus subtilis o	nge 177, abstract no. Ohio, US; L.I. ening genes of the nesis operon of the plasmid pBR322 nia coli", & GENETIKA	1-5,8-13,15	C 12 P
		-/-		
	The present search report has I	Date of completion of the search	<u> </u>	Examiner
THE	HAGUE	28-09-1990	DES	CAMPS J.A.
X : par Y : par doc	CATEGORY OF CITED DOCUME ticularly relevant if taken alone ticularly relevant if combined with an ament of the same category anological background	E : earlier pate after the fli other D : document of	rinciple underlying the at document, but publing date cited in the applicationted for other reasons	dished on, or

EPO FORM 1503 03.82 (P0401)





# **EUROPEAN SEARCH REPORT**

Application Number

EP 90 11 1916

### Relevant to claim APPLICATION OF THE APPLICATION (Int. Cl.5)  15, 1-5,8- 13,15  19, 1-5,8- 13,15  10. V. Ichia S S S S S S S S S S S S S S S S S S S
19, 1-5,8- 13,15  19, 1-5,8- 13,15  1-5,8- 13,15  23, 1-5,8- 13,15  TECHNICAL FIELDS SEARCHED (Int. CL.5)
ract   13,15
t no. 13,15  TECHNICAL FIELDS SEARCHED (Int. CL5)
ims
990 DESCAMPS J.A.
: theory or principle underlying the invention : earlier patent document, but published on, or after the filing date